

09/475704

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2003, 14:02:05 ; Search time 1380 Seconds
(without alignments)
1778.679 Million coli updates/sec

Title: US-09-475-704A-1
Perfect score: 60
Sequence: 1 gacatcaagcagggccccc...tggacggctcttcaagacc 60

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 1520254

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb_ba:
2: gb_btq:
3: gb_in:
4: gb_om:
5: gb_ov:
6: gb_pat:
7: gb_ph:
8: gb_pl:
9: gb_pr:
10: gb_ro:
11: gb_sts:
12: gb_sy:
13: gb_un:
14: gb_vl:
15: em_ba:
16: em_fun:
17: em_hum:
18: em_in:
19: em_ma:
20: em_om:
21: em_or:
22: em_ov:
23: em_pat:
24: em_ph:
25: em_pl:
26: em_ro:
27: em_sts:
28: em_un:
29: em_vl:
30: em_htg_hum:
31: em_htg_inv:
32: em_htg_other:
33: em_htg_mus:
34: em_htg_pln:
35: em_htg_rod:
36: em_htg_mam:
37: em_htg_vrt:
38: em_sy:
39: em_htgo_hum:
40: em_htgo_mus:
41: em_htgo_other:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	60	100.0	60	6	AX455885	AX455885 Sequence
2	60	100.0	60	6	AX455937	AX455937 Sequence
3	55.2	92.0	60	6	AX455886	AX455886 Sequence
4	37.6	62.7	60	6	AX455910	AX455910 Sequence
5	36	60.0	60	6	AX455938	AX455938 Sequence
6	21.2	35.3	91	11	AF275608	AF275608 Bos tauru
7	20.4	34.0	60	6	AR169921	AR169921 Sequence
8	20	33.3	81	6	AX480659	AX480659 Sequence
9	19.8	33.0	97	1	AF223423	AF223423 Pseudomon
10	19.6	32.7	84	9	HSU91013	U91013 Homo sapien
11	19.4	32.3	78	6	AX480660	AX480660 Sequence
12	19.2	32.0	51	6	AX158488	AX158488 Sequence
13	19.2	32.0	97	6	AX088801	AX088801 Sequence
14	19	31.7	53	6	108503	108503 Sequence 17
15	19	31.7	65	6	108502	108502 Sequence 16
16	18.8	31.3	30	6	AX026570	AX026570 Sequence
17	18.8	31.3	64	6	AX453567	AX453567 Sequence
18	18.8	31.3	74	5	AY227159	AY227159 Alca tord
19	18.6	31.0	47	6	AX195033	AX195033 Sequence
20	18.6	31.0	51	6	AX163045	AX163045 Sequence
21	18.4	30.7	51	9	HSDNASPAT	222386 H.sapiens D
22	18.4	30.7	62	6	AL0076	AL0076 Epstein-Bar
23	18.4	30.7	62	14	FBVCLGAG	X07531 EBV B95-8 C
24	18.4	30.7	69	10	AY205871	AY205871 Mus muscu
25	18.4	30.7	74	6	AX098157	AX098157 Sequence
26	18	30.0	80	14	AF040847	AF040847 Hepatitis
27	18	30.0	80	14	AF040849	AF040849 Hepatitis
28	18	30.0	98	3	ASBP23A2	AF134278 Agrotis s
29	18	30.0	98	3	ASBP23B2	AF134280 Agrotis s
30	18	30.0	98	3	ASBP24A2	AF134282 Agrotis s
31	18	30.0	98	3	ASBP25A2	AF134284 Agrotis s
32	18	30.0	98	3	ASBP26A2	AF134286 Agrotis s
33	18	30.0	98	3	ASBP27A2	AF134288 Agrotis s
34	18	30.0	100	6	AR078713	ARC78713 Sequence
35	18	30.0	100	6	AR198693	AR198693 Sequence
36	18	30.0	100	6	125137	125137 Sequence 33
37	17.8	29.7	40	6	AR142031	AR142031 Sequence
38	17.8	29.7	40	6	159898	159898 Sequence 25
39	17.8	29.7	40	6	168065	168065 Sequence 3
40	17.8	29.7	40	6	186756	186756 Sequence 25
41	17.8	29.7	40	6	195781	195781 Sequence 25
42	17.8	29.7	47	6	AX194662	AX194662 Sequence
43	17.8	29.7	51	9	578515	578515 type I kera
44	17.8	29.7	56	6	AX103647	AX103647 Sequence
45	17.8	29.7	63	6	AR310621	AR310621 Sequence

ALIGNMENTS

RESULT 1
AX455885
LOCUS AX455885
DEFINITION Sequence 1 from Patent WO0204493.
ACCESSION AX455885
VERSION AX455885.1 GI:21714878
KEYWORDS
SOURCE Human immunodeficiency virus
ORGANISM Human immunodeficiency virus
Viruses; Retroviral viruses; Retroviridae; Lentivirus; Primate
REFERENCE
AUTHORS zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.
TITLE Polynucleotides encoding antigenic hiv type c polypeptides,
polypeptides and uses thereof

60 bp DNA Linear FAT 06-JUL-2002


```
RESULT 13
AX088801
LOCUS AX088801 97 bp DNA linear PAT 17-MAR-2001
DEFINITION Sequence 127 from Patent WO0114416.
ACCESSION AX088801
VERSION AX088801.1 GI:13397597
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
          artificial sequences.
REFERENCE
AUTHORS Neuper,M.P., McClements,W.L., Jansen,K.U., Schuitz,L.D., Chen,L.
          and Wang,X.M.
TITLE Synthetic human papillomavirus genes
JOURNAL Patent: WO 0114416-A 127 01-MAR-2001;
          Merck & Co., Inc. (US)
FEATURES
          Location/Qualifiers
          source
            1..97
              /organism="synthetic construct"
              /mol_type="genomic DNA"
              /db_xref="taxon:32630"
              /note="Codon-Optimized HPV16 E2 fragment"
BASE COUNT 29 a 30 c 24 g 14 t
ORIGIN
Query Match 32.0%; Score 19.2; DB 6; Length 97;
Best Local Similarity 67.5%; Pred. No. 2.2e+05;
Matches 27; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 3 CATCAGCAGGGCCCAAGAGAGCCCTCCGGCGACTACGTTG 42
   |||||  |||||  |||||  |||||  |||||  |||||
Db 18 CACCATGCAGACTCCCAAGAGAGAGCCCTCCGGCGAGCGGCTG 57

RESULT 14
I08503
LOCUS I08503 53 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 17 from Patent WO 8701728.
ACCESSION I08503
VERSION I08503.1 GI:588795
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 53)
AUTHORS Fiddes,J.C. and Abraham,J.A.
TITLE RECOMBINANT FIBROBLAST GROWTH FACTORS
JOURNAL Patent: WO 8701728-A 17 26-MAR-1987;
          Location/Qualifiers
          source
            1..53
              /organism="unknown"
BASE COUNT 11 a 17 c 15 g 10 t
ORIGIN
Query Match 31.7%; Score 19; DB 6; Length 53;
Best Local Similarity 65.1%; Pred. No. 2.7e+05;
Matches 28; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 11 AGGGCCCCAAGAGAGCCCTCCGGCGACTACGTGGACCGCTTCTT 53
   ||| |||||  ||  |||||  |||||  |||||  |||||
Db 8 AGGACCCCAAGCGGCTGTACTGCAAGAACGGGGGCTTCTTCT 50

RESULT 15
I08502
LOCUS I08502 65 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 16 from Patent WO 8701728.
ACCESSION I08502
VERSION I08502.1 GI:588794
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
```

```
Unclassified.
REFERENCE 1 (bases 1 to 65)
AUTHORS Fiddes,J.C. and Abraham,J.A.
TITLE RECOMBINANT FIBROBLAST GROWTH FACTORS
JOURNAL Patent: WO 8701728-A 16 26-MAR-1987;
          Location/Qualifiers
          source
            1..65
              /organism="unknown"
BASE COUNT 11 a 23 c 19 g 12 t
ORIGIN
Query Match 31.7%; Score 19; DB 6; Length 65;
Best Local Similarity 65.1%; Pred. No. 2.6e+05;
Matches 28; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 11 AGGGCCCCAAGAGAGCCCTCCGGCGACTACGTGGACCGCTTCTT 53
   ||| |||||  ||  |||||  |||||  |||||  |||||
Db 20 AGGACCCCAAGCGGCTGTACTGCAAGAACGGGGGCTTCTTCT 62

Search completed: September 22, 2003, 14:56:51
Job time : 1386 secs
```


Db 61 CTGCGCCCGCGGCGCAAGAGACACTACATGCTGAGCAGCTGGTGTGGGCCAGCCGTGAG 120
QY 121 CTGGAGGGGTTCGCCCTGAAACCCCGGCTGCTGGAGACCGCGGAGGGTGCAGCAGATC 180
Db 121 CTGGAGGGGTTCGCCCTGAAACCCCGGCTGCTGGAGACCGCGGAGGGTGCAGCAGATC 180
QY 181 ATGAAGCAGGTGCAGCCCGGCTGCAGACCGGCGACACAGAGAGCTGGAGGCCCTGACAA 240
Db 181 ATGAAGCAGGTGCAGCCCGGCTGCAGACCGGCGACACAGAGAGCTGGAGGCCCTGACAA 240
QY 241 ACCGTGGCCACCCCTGTACTGGGTGCACCGCGGCTATCGAGGTCCGGGACACCAAGAGAGCC 300
Db 241 ACCGTGGCCACCCCTGTACTGGGTGCACCGCGGCTATCGAGGTCCGGGACACCAAGAGAGCC 300
QY 301 CTGGACACATCGAGGAGGAGCAGACACAGTCCAGCAGAGAACCCAGCAGGCGCCAGAGAG 360
Db 301 CTGGACACATCGAGGAGGAGCAGACACAGTCCAGCAGAGAACCCAGCAGGCGCCAGAGAG 360
QY 361 GCCGACGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGGCCAGATGGTG 420
Db 361 GCCGACGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGGCCAGATGGTG 420
QY 421 CACCAGGCCATCAGCCCGCCGACCCCTGAGACCCCTGGGTGAAGGTGATGAGGAGAGAGCC 480
Db 421 CACCAGGCCATCAGCCCGCCGACCCCTGAGACCCCTGGGTGAAGGTGATGAGGAGAGAGCC 480
QY 481 TTACGCCCGGAGGTGATCCCATGTTCCACCGCCCTGAGCAGGCGGCGACCCCGCAGAGAG 540
Db 481 TTACGCCCGGAGGTGATCCCATGTTCCACCGCCCTGAGCAGGCGGCGACCCCGCAGAGAG 540
QY 541 CTGAACACGATGTTGAACACCGTGGGCGGCCACACAGGCGGCCATGCAGATGCTGAAGGAC 600
Db 541 CTGAACACGATGTTGAACACCGTGGGCGGCCACACAGGCGGCCATGCAGATGCTGAAGGAC 600
QY 601 ACCATCAACGAGGAGGCGCGGAGTGGGACCGCCTGCACCCCGTGCAGGCGCGGCCCTGTG 660
Db 601 ACCATCAACGAGGAGGCGCGGAGTGGGACCGCCTGCACCCCGTGCAGGCGCGGCCCTGTG 660
QY 661 GCCCGCGGCAGATGCGGCGACCCCGCGGAGCGACATCGCGGCGGCGCACCGACACCGTG 720
Db 661 GCCCGCGGCAGATGCGGCGACCCCGCGGAGCGACATCGCGGCGGCGCACCGACACCGTG 720
QY 721 CAGGAGCAGATCGCCTGGATGACCAACACCCCGCGGCGGCGGCGGCGACATCTACAAG 780
Db 721 CAGGAGCAGATCGCCTGGATGACCAACACCCCGCGGCGGCGGCGGCGGCGGCGGCGGCG 780
QY 781 CGGTGGATCATCTCGGCTGAACAAGATCGTGGGATGTACAGCCCGGTGAGGATCTG 840
Db 781 CGGTGGATCATCTCGGCTGAACAAGATCGTGGGATGTACAGCCCGGTGAGGATCTG 840
QY 841 GACATCCCGCAGGGCGCCCAAGAGGAGCCCTTCGCGGACTACCTGGACCGCTTCTTCAAGAC 900
Db 841 GACATCCCGCAGGGCGCCCAAGAGGAGCCCTTCGCGGACTACCTGGACCGCTTCTTCAAGAC 900
QY 901 CTGCGCGCGCAGCAGGCGCCAGCAGGAGCGTGAAGAAGTGGATGACCGACCCCTGCTGCTG 960
Db 901 CTGCGCGCGCAGCAGGCGCCAGCAGGAGCGTGAAGAAGTGGATGACCGACCCCTGCTGCTG 960
QY 961 CAGAACGCCAACCCCGACTGCAAGACCATCTGCGGCTGTCTGGCGCGCGGCGGCGGCGGCG 1020
Db 961 CAGAACGCCAACCCCGACTGCAAGACCATCTGCGGCTGTCTGGCGCGCGGCGGCGGCGGCG 1020
QY 1021 GAGGAGATGTACCGCTGCGCAGGCGGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1080
Db 1021 GAGGAGATGTACCGCTGCGCAGGCGGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1080
QY 1081 GCGGAGGCGATGAGCGGCGCAACAGCGTGANCACTCATATGAGAGAGCAACTTCAAG 1140
Db 1081 GCGGAGGCGATGAGCGGCGCAACAGCGTGANCACTCATATGAGAGAGCAACTTCAAG 1140
QY 1141 GCGCCCGCGGCGCAACGTCAAGTCTTCAACTCGCGCAAGAGGCGGCGGCGGCGGCGGCGG 1200
Db 1141 GCGCCCGCGGCGCAACGTCAAGTCTTCAACTCGCGCAAGAGGCGGCGGCGGCGGCGGCGG 1200

QY 1201 TGCCGCGCCCGCGCAGAGAGGCTGCTGGAAGTGGGCAAGGAGGGGCCACCATGTAAG 1260
Db 1201 TGCCGCGCCCGCGCAGAGAGGCTGCTGGAAGTGGGCAAGGAGGGGCCACCATGTAAG 1260
QY 1261 GACTGCACCGAGCGCCAGGCGCAACTTCCTGGCAAGATCTGGCCCGCAGCCACAAGGGCCG 1320
Db 1261 GACTGCACCGAGCGCCAGGCGCAACTTCCTGGCAAGATCTGGCCCGCAGCCACAAGGGCCG 1320
QY 1321 CCGGGAACCTTCCTGCAGAACCGCAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
Db 1321 CCGGGAACCTTCCTGCAGAACCGCAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
QY 1381 CCGGCGGAGAGCTTCGCTTCGAGGAGACACCCCGCGCGCGCGCGCGCGCGCGCGCGCG 1440
Db 1381 CCGGCGGAGAGCTTCGCTTCGAGGAGACACCCCGCGCGCGCGCGCGCGCGCGCGCGCG 1440
QY 1441 CCGGAGCCCTACCGCGGAGCGCGCTGACCGCGCTGCGCAGCGCTTCTGGCAGCGCGCGCG 1500
Db 1441 CCGGAGCCCTACCGCGGAGCGCGCTGACCGCGCTGCGCAGCGCTTCTGGCAGCGCGCGCG 1500
QY 1501 AGCCAGTAA 1509
Db 1501 AGCCAGTAA 1509

RESULT 2

US-09-967-464-68
; Sequence 68, Application US/09967464
; Publication No. US20030138453A1
; GENERAL INFORMATION:
; APPLICANT: O'Hagan, Derek
; APPLICANT: Otten, Gillis
; APPLICANT: Donnelly, John J.
; APPLICANT: Polo, John M.
; APPLICANT: Barnett, Susan
; APPLICANT: Singh, Mamohan
; APPLICANT: Glmer, Jeffrey
; APPLICANT: Dubensky, Jr., Thomas W.
; TITLE OF INVENTION: MICROPARTICLES FOR DELIVERY OF HETEROLOGOUS NUCLEIC ACIDS
; FILE REFERENCE: PP16269.004
; CURRENT APPLICATION NUMBER: US/09/967,464
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 60/236,105
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 60/315,905
; PRIOR FILING DATE: 2001-08-30
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 68
; LENGTH: 1509
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
US-09-967-464-68

Query Match 99.7%; Score 1504.2; DB 12; Length 1509;
Best Local Similarity 99.2%; Pred. No. 8.5e-299;
Matches 1497; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGCGCCCGCGCCAGCACTCTGCGCGGCGGAGAGCTGACAAAGTGGGAGAGATCCGC 60
Db 1 ATGGGCGCCCGCGCCAGCACTCTGCGCGGCGGAGAGCTGACAAAGTGGGAGAGATCCGC 60
QY 61 CTGGCGCCCGCGCGGCGGCAAGCACTACATCTGAGCACCTGGTGTGGGCGAGCCGCGAG 120
Db 61 CTGGCGCCCGCGCGGCGGCAAGCACTACATCTGAGCACCTGGTGTGGGCGAGCCGCGAG 120
QY 121 CTGGAGGGCTTCGCCCTGACCCCGGCTGTGGAGACCGCGGAGGCTGCAAGCAGATC 180
Db 121 CTGGAGGGCTTCGCCCTGACCCCGGCTGTGGAGACCGCGGAGGCTGCAAGCAGATC 180
QY 181 ATGAAGCAGCTGCAGCCCGCGCTGCAGACCGGCGACCGAGAGCTGCGCAGGCTGTACAA 240
Db 181 ATGAAGCAGCTGCAGCCCGCGCTGCAGACCGGCGACCGAGAGCTGCGCAGGCTGTACAA 240

Db 181 ATCAAGCGCTCGAGCCCGCCCTGCAGACCGGACCGAGAGCTCGGACGCTGTACAAC 240

QY 241 ACCGIGGCCACCCCTGTACTCGTCCACGCCCGCATCGAGGTCCGCGACACCAAGAGGCC 300

Db 241 ACCGTGGCCACCCCTGTACTCGTCCACGCCCGCATCGAGGTSCGCGACACCAAGAGGCC 300

QY 301 CTGGACAAGATCGAGGAGGAGGAGACAAGTCCACGACGAGACCCACGAGGCCCAAGGAG 360

Db 301 CTGGACAAGATCGAGGAGGAGGAGACAAGTCCACGACGAGACCCACGAGGCCCAAGGAG 360

QY 361 GCCGACGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACTGCAGGGCCAGATGGTG 420

Db 361 GCCGACGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACTGCAGGGCCAGATGGTG 420

QY 421 CACGAGGCCATCAGCCCCCGCACCCCTGAACGGCTGGTGAAGTGATCGAGGAGAGGCC 480

Db 421 CACGAGGCCATCAGCCCCCGCACCCCTGAACGGCTGGTGAAGTGATCGAGGAGAGGCC 480

QY 481 TTCAGCCCCGAGGTGATCCCCATGTTCCACCGCCCTGAGCGAGGGGCCACCCCCAGGAC 540

Db 481 TTCAGCCCCGAGGTGATCCCCATGTTCCACCGCCCTGAGCGAGGGGCCACCCCCAGGAC 540

QY 541 CTGAACACGATGTTGAACACCGTGGGGCGGCCACGAGCCGCCATGCAGATGCTGAAGGAC 600

Db 541 CTGAACACGATGTTGAACACCGTGGGGCGGCCACGAGCCGCCATGCAGATGCTGAAGGAC 600

QY 601 ACCATCAACGAGGAGGCCGCCGAGTGGGACCGCTGCACCCCTGCAGGGGGCCCGGTG 660

Db 601 ACCATCAACGAGGAGGCCGCCGAGTGGGACCGCTGCACCCCTGCAGGGGGCCCGGTG 660

QY 661 GCCCGCGCCAGATGCGGCGACCCCGCGGCAAGCACTGCGCGGCGGCCACGAGCCCTG 720

Db 661 GCCCGCGCCAGATGCGGCGACCCCGCGGCAAGCACTGCGCGGCGGCCACGAGCCCTG 720

QY 721 CAGGAGCAGATCGCTGGATGACGAGCAACCCCGCCGTGCCGGTGGGGGACATCAAG 780

Db 721 CAGGAGCAGATCGCTGGATGACGAGCAACCCCGCCGTGCCGGTGGGGGACATCAAG 780

QY 781 CGGTGGATCATCTGGGCTTGAACAAGATCGTGGGATGTACAGCCCCGTGAGCATCTG 840

Db 781 CGGTGGATCATCTGGGCTTGAACAAGATCGTGGGATGTACAGCCCCGTGAGCATCTG 840

QY 841 GACATCCGCGAGGCCCGCAAGGAGCCCTTCCGGCACTACGIGGACCGCTTCTCAAGAC 900

Db 841 GACATCCGCGAGGCCCGCAAGGAGCCCTTCCGGCACTACGIGGACCGCTTCTCAAGAC 900

QY 901 CTGCGCGCGGAGGAGGCCCGCAGGACGTTGAAGACTGGATGACCGAGACCCCTGCTGTG 960

Db 901 CTGCGCGCGGAGGAGGCCCGCAGGACGTTGAAGACTGGATGACCGAGACCCCTGCTGTG 960

QY 961 CAGAACGCCAACCCTGACTGCAGACCATCTCGCGGCTCTCGCGCCCGGCGCACCCCTG 1020

Db 961 CAGAACGCCAACCCTGACTGCAGACCATCTCGCGGCTCTCGCGCCCGGCGCACCCCTG 1020

QY 1021 GAGGAGATGATGACCGCTGCCAGGGCGTGGGGCGCCCGGCCACAAAGGCCCGTCTG 1080

Db 1021 GAGGAGATGATGACCGCTGCCAGGGCGTGGGGCGCCCGGCCACAAAGGCCCGTCTG 1080

QY 1081 GCCGAGGATGAGCCAGGCCACAGCGTGAACATCATGATGAGAAAGAGCAACTTCAAG 1140

Db 1081 GCCGAGGATGAGCCAGGCCACAGCGTGAACATCATGATGAGAAAGAGCAACTTCAAG 1140

QY 1141 GGCCCGCGCGCAAGTCAAGTCTCAACTGGCGCAAGGAGGCCACATCGCAAGAAG 1200

Db 1141 GGCCCGCGCGCAAGTCAAGTCTCAACTGGCGCAAGGAGGCCACATCGCAAGAAG 1200

QY 1201 TGCCCGCCCCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGGAGGGCCACCATGAAG 1260

Db 1201 TGCCCGCCCCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGGAGGGCCACCATGAAG 1260

QY 1261 GACTGCACCGAGCGCAGGCCAACTTCTTGGCAGATCTGGCCAGCCACAGGGCCCG 1320

Db 1261 GACTGCACCGAGCGCAGGCCAACTTCTTGGCAGATCTGGCCAGCCACAGGGCCCG 1320

QY 1321 CCCGGCAACTTCTGCAGAACCGCAGCGAGCCGCGGCCCCACCCCTGCCACCGCCCC 1380

Db 1321 CCCGGCAACTTCTGCAGAACCGCAGCGAGCCGCGGCCCCACCCCTGCCACCGCCCC 1380

QY 1381 CCCGCGAGAGCTTCCGCTTCGAGGAGACCAACCCCGCCCGCCCAAGGAGGCCAAGGAC 1440

Db 1381 CCCGCGAGAGCTTCCGCTTCGAGGAGACCAACCCCGCCCGCCCAAGGAGGCCAAGGAC 1440

QY 1441 CGCGAGCCCTACCGCGAGCCCTGACCGCCCTGCGCAGCCCTGTTGCGCAGCGGCCCTG 1500

Db 1441 CGCGAGCCCTACCGCGAGCCCTGACCGCCCTGCGCAGCCCTGTTGCGCAGCGGCCCTG 1500

QY 1501 AGCCAGTAA 1509

Db 1501 AGCCAGTAA 1509

RESULT 3

US-09-967-464-63

; Sequence 63, Application US/09967464

; Publication No. US20030138453A1

GENERAL INFORMATION:

; APPLICANT: O'Hagan, Derek

; APPLICANT: Otten, Gillis

; APPLICANT: Donnelly, John J.

; APPLICANT: Polo, John M.

; APPLICANT: Barnett, Susan

; APPLICANT: Singh, Mamohan

; APPLICANT: Ulmer, Jeffrey

; APPLICANT: Dubensky, Jr., Thomas W.

; TITLE OF INVENTION: MICROPARTICLES FOR DELIVERY OF HETEROLOGOUS NUCLEIC ACIDS

; FILE REFERENCE: PPI6269.004

; CURRENT APPLICATION NUMBER: US/09/967,464

; CURRENT FILING DATE: 2002-04-11

; PRIOR APPLICATION NUMBER: 60/236,105

; PRIOR FILING DATE: 2000-09-28

; PRIOR APPLICATION NUMBER: 60/315,905

; PRIOR FILING DATE: 2001-08-30

; NUMBER OF SEQ ID NOS: 68

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 63

; LENGTH: 1479

; TYPE: DNA

; ORGANISM: Human immunodeficiency virus type 1

US-09-967-464-63

Query Match 84.6%; Score 1276.8; DB 12; Length 1479;

Best Local Similarity 92.5%; Pred. No. 2.7e-252;

Matches 1399; Conservative 0; Mismatches 77; Indels 36; Gaps 4;

QY 1 ATGGGCGCGCGCCAGCATCTCGCGCGGAGAAAGTGGAGAGATCGGC 60

Db 1 ATGGGCGCGCGCCAGCATCTCGCGCGGCGGCAAGCTGGACGCTGGGAGCGCATCCGC 60

QY 61 CTGCGCCCGCGCGCAAGCAACTACATGCTGAAGCACTGCTGGAGCCCTGGTGTGGCCAGCGCGAG 120

Db 61 CTGCGCCCGCGCGCAAGCAAGTGTCTACATGATGAAGCACTGCTGTGGGCCAGCGCGAG 120

QY 121 CTGAGGGCTTCGCCCTGAACCCCGCGCTGCTGGAGACCGCGGAGGCTGCAAGCAGATC 180

Db 121 CTGAGAAAGTTCGCCCTGAACCCCGCGCTGCTGGAGACCGCGGAGGCTGCAAGCAGATC 180

QY 181 ATGAAGCAGCTGCAGCCCGCCCTGCAGACCGCGCACCCGAGGAGCTGCCAGCCTGTACAAC 240

Db 181 ATCCGCCAGCTGCACCCCGCCCTGCAGACCGCGCACCGAGGAGCTGAAGAGCCTGTTCAAC 240

QY 241 ACCGTGGCCACCCTGTACTGCTGCAGCGCGGCATCGAGGTCCGCGACACCAAGGAGGCC 300

Db 241 ACCGTGGCCACCCTGTACTGCTGCAGCAGAAAGATCGAGTCCGCGACCAAGGAGGCC 300

QY 301 CTGGACAAGATCGAGGAGGAGCAGAACAACTCCAGCAAGACCCAGAGCCAGGCCAAGGAG 360

Db 301 CTGGACAAAGATCGAGGAGGACGAGAAAGTGCCACGAGAAGATCCAGCAUNCGAGGCC 360
Qy 361 GCCGAC---GGCAAGGTGAGCCAGAGAACTACCCATCGTIGCAGAACTCTGAGGGCCAGATG 417
Db 361 GCCGACAAAGGCAAGGTGAGCCAGAGAACTACCCATCGTIGCAGAACTCTGAGGGCCAGATG 420
Qy 418 GTGACACAGGCCATACGCCCCCGGACCCCTGAACGCTGGTGAAGGTGATCGAGGAGAAG 477
Db 421 GTGACACAGGCCATACGCCCCCGGACCCCTGAACGCTGGTGAAGGTGATCGAGGAGAAG 480
Qy 478 GCCTTCAGCCCGAGGTGATCCCATGTTTACCGGCTGAGCGAGGGCGCCACCCGCCAG 537
Db 481 GCCTTCAGCCCGAGGTGATCCCATGTTTACCGGCTGAGCGAGGGCGCCACCCGCCAG 540
Qy 538 GACCTGACACAGATGTTGAACACCGTGGCGGGCCACGAGGGCCGATGAGATGCTGAAG 597
Db 541 GACCTGACACAGATGTTGAACACCGTGGCGGGCCACGAGGGCCGATGAGATGCTGAAG 600
Qy 598 GACACCATCAACGAGGAGGCCCGGAGTGGGACCGCTGCACCCGCTGAGGCGCGGCCGCC 657
Db 601 GACACCATCAACGAGGAGGCCCGGAGTGGGACCGCTGCACCCGCTGAGGCGCGGCCGCC 660
Qy 658 GTGGCCCGCGCCAGATCGCGGACCCCGCGGCGGACGACATCGCGGCGCCACGACGACC 717
Db 661 ATCGCCCGCGCCAGATCGCGGACCCCGCGGCGGACGACATCGCGGCGCCACGACGACC 720
Qy 718 CTGACGAGCAGATCGCTCGATGACGACCAACCCCGCGCTGCCCTGGGCGGACATCTAC 777
Db 721 CTGACGAGCAGATCGCTCGATGACGACCAACCCCGCGCTGCCCTGGGCGGACATCTAC 780
Qy 778 AAGCGGTGGATCATCTCGGCTGAACAAAGATCGTGGGATGTACAGCCCGCTGAGCATC 837
Db 781 AAGCGGTGGATCATCTCGGCTGAACAAAGATCGTGGGATGTACAGCCCGCTGAGCATC 840
Qy 838 CTGGACATCGGCCAGGSCCCCAAGGAGCCCTTCCGCGACIACGTGGACCGCTTCTCAAG 897
Db 841 CTGGACATCAAGCAGGSCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTCAAG 900
Qy 898 ACCCTGCGCGCGAGCAGGCCACCCAGGACCGTGAAGAACTGGATGACCGGAGACCGCTGTG 957
Db 901 ACCCTGCGCGCGAGCAGGCCACCCAGGAGGTGAAGAACTGGATGACCGGAGACACCGCTGTG 960
Qy 958 GTGCAGAAAGCCCAACCCGCACTGCAAGACCATCTCGCGGCTCTCGGCGCCCGGCGCAC 1017
Db 961 GTGCAGAAAGCCCAACCCGCACTGCAAGACCATCTCGCGGCTCTCGGCGCCCGGCGCAC 1020
Qy 1018 CTGGAGGAGATGATGACCGGCTGCCAGGCGGTGGGCGCCCGCGGCAAGGCGCGGCTG 1077
Db 1021 CTGGAGGAGATGATGACCGGCTGCCAGGCGGTGGGCGCCCGCGGCAAGGCGCGGCTG 1080
Qy 1078 CTGGCGGAGCGGATGAGCGAGGCCCAACACGCGTGAACATCATGATGCAAGAGCAACTC 1137
Db 1081 CTGGCGGAGCGGATGAGCGAGGCCCAACACG---AGCGTGATGATGCAAGAGCAACTC 1137
Qy 1138 AAGGGCCCGCGCGCAACGTCAAGTGTCTCAACTGGGCAAGGAGGGCCACATCGGCAAG 1197
Db 1138 AAGGGCCCGCGCGCAACGTCTCAAGTGTCTCAACTGGGCAAGGAGGGCCACATCGGCGC 1197
Qy 1198 AACTGCCGCGCCCGCGCAAGAGGGCTGCTGGAAGTGGGCAAGGAGGGCCACCCAGATG 1257
Db 1198 AACTGCCGCGCCCGCGCAAGAGGGCTGCTGGAAGTGGGCAAGGAGGGCCACCCAGATG 1257
Qy 1258 AAGGACTGCACCGGAGCGGCGAGGCCAACTTCCTGGGCAAGATCTGGCGGAGGCCACAGGGC 1317
Db 1258 AAGGACTGCACCGGAGCGGCGAGGCCAACTTCCTGGGCAAGATCTGGCGGAGGCCACAGGGC 1317
Qy 1318 CGCCCGGCAACTTCCTGCAAGAACCGGACGAGCGCGCGCCCGCCACCGTGGCCATCGCC 1377
Db 1318 CGCCCGGCAACTTCCTGCAAGACCG-----CGCCGAGCCCGCC 1389
Qy 1376 CGCCCGGCGAGAGCTTCGCTTCGAGGAGACACCCCGCCCGCCCAAGCAGGAGGCGAAG 1437
Db 1360 CGCCCGGCGAGAGCTTCGCTTCGAGGAGACACCCCGCGCCCGGCGGAGAGGAGGCAAG 1419

Qy 1438 GACCGCGAGCCCTACCGGAGCGCCCTGACCGGCTGACCGGCTGTTCGGCAGCGGCCCC 1497
Db 1420 G-----ACCGGAGACCCCTGACCGGCTGAAGAGCCTGTTCGGCAGCAGCCCC 1457
Qy 1498 CTGAGCCAGTAA 1509
Db 1468 CTGAGCCAGTAA 1479
RESULT 4
US-10-190-435-18
; Sequence 18, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES, AND USES THEREOF
; FILE REFERENCE: PPI8133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190.435
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 3162
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GagRTmul_C
US-10-190-435-18

Query Match 84.5%; Score 1275.2; DB 12; Length 3162;
Best Local Similarity 92.5%; Pred. NO. 5.5e-252;
Matches 1398; Conservative 0; Mismatches 78; Indels 36; Gaps 4;

Qy 1 ATGGGGCGCCCGCCAGCATCTCTGCGGCGGCAAGACTGAGCAAGTGGGAGAGATCGCG 60
Db 7 ATGGGGCGCCCGCCAGCATCTCTGCGGCGGCAAGACTGAGCAAGTGGGAGAGATCGCG 66
Qy 61 CTGCGCCCGCGGCGGCAAGACTGAGCAAGTGGGAGAGACTGAGCAAGTGGGAGAGATC 120
Db 67 CTGCGCCCGCGGCGGCAAGACTGAGCAAGTGGGAGAGACTGAGCAAGTGGGAGAGATC 126
Qy 121 CTGAGGAGCTGCGCCCTGAGACCCCGGCTGCTGAGAGAGAGAGAGAGAGAGAGATC 180
Db 127 CTGAGAGAGTTCGCTGAGACCCCGGCTGCTGAGAGAGAGAGAGAGAGAGATC 186
Qy 181 ATGAAGCAGCTGCAGCCCGGCTGCAGACCGGCGGAGAGAGAGAGAGAGAGATC 240
Db 187 ATCGCGCAGCTGCAGCCCGGCTGCAGACCGGCGGAGAGAGAGAGAGAGAGATC 246
Qy 241 ACCGTGGCCACCCCTGTACTGGTGCACCGCGGCTGAGAGAGAGAGAGAGAGAGAG 300
Db 247 ACCGTGGCCACCCCTGTACTGGTGCACCGAGAGAGATCGAGGTCCCGAGAGAGAG 306
Qy 301 CTGGACAAGATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
Db 307 CTGGACAAGATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 366
Qy 361 GCGGAC---GCGAAGGTGAGCCAGAACTACCCCATCTGTCAGAGAACTGAGGAGGAG 417
Db 367 GCGGACAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 426
Qy 418 GTGCACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 477
Db 427 GTGCACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 486
Qy 478 GCCTTCAGCCCGAGGTGATCCCATGTTTACCGGCTGAGCGGAGGGCGCCACCCCGAG 537
Db 478 GCCTTCAGCCCGAGGTGATCCCATGTTTACCGGCTGAGCGGAGGGCGCCACCCCGAG 537

```
Db 487 GCCTTCAGCCCGGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGCGCGCCACGCCGCCAS 546
QY 538 GACCTGAACACAGATGTTGAACACCGTGGGCGGCGACCGAGCCGCCATGACAGTCTGTAAG 597
Db 547 GACCTGAACACAGATGTTGAACACCGTGGGCGGCGACCGAGCCGCCATGACAGTCTGTAAG 606
QY 598 GACACCATCAACGAGGAGGCGCGGAGTGGGACCGCTGCACCCCGTGCAGGCGCGGCCAC 657
Db 607 GACACCATCAACGAGGAGGCGCGGAGTGGGACCGCTGCACCCCGTGCAGGCGCGGCCAC 666
QY 658 GTGGCCCCCGCCAGATGCGCGACCCCGCGGACCGGACATCGCGCGCGCGCCACGACAC 717
Db 667 ATCGCCCCCGCCAGATGCGCGAGCCCGCGGACCGGACATCGCGCGCGCGCCACGACAC 726
QY 718 CTGCAGGAGCAGATCGCCCTGGATGACCGACCAACCCCGCGTGCCTGGGCGACATCTAC 777
Db 727 CTGCAGGAGCAGATCGCCCTGGATGACCGACCAACCCCGCGTGCCTGGGCGACATCTAC 786
QY 778 AAGCGGTGGATCATCTCTGGGCTGAACAAGATCGTGGGATGTACAGCCCGCTGAGCATC 837
Db 787 AAGCGGTGGATCATCTCTGGGCTGAACAAGATCGTGGGATGTACAGCCCGCTGAGCATC 846
QY 838 CTGCACATCCGCGAGGCGCCCAAGAGCCCTTCGCGGACIACGTGGACCGCTCTCTCAAG 897
Db 847 CTGCACATCAAGCAGGCGCCCAAGAGCCCTTCGCGGACIACGTGGACCGCTCTCTCAAG 906
QY 898 ACCCTGCGCGCGGAGGCGCCACCGAGGACGTGAAGAACTGGATGACCGAGACCTGCTG 957
Db 907 ACCCTGCGCGCGGAGGCGCCACCGAGGACGTGAAGAACTGGATGACCGAGACCTGCTG 966
QY 958 GTGCAGAACCCCAAGCCCGACTGCAAGACCATCTCTGCGCGCTCTCGGCCCGCGGCCACC 1027
Db 967 GTGCAGAACCCCAAGCCCGACTGCAAGACCATCTCTGCGCGCTCTCGGCCCGCGGCCACC 1026
QY 1018 CTGAGGAGATGATGACCCCGCTGCCAGCGCGTGGGCGCGCCCGGCGGACAAAGCGCCGCTG 1077
Db 1027 CTGAGGAGATGATGACCCCGCTGCCAGCGCGTGGGCGCGCCCGGCGGACAAAGCGCCGCTG 1086
QY 1078 CTGGCGGAGGCGATGAGCCAGGCGCCACAGCGGTGACATCATGATGACAGAGACACTTC 1137
Db 1087 CTGGCGGAGGCGATGAGCCAGGCGCCACACACC--AGCGTATGATGACAGAGAGCACTTC 1143
QY 1138 AAGGCCCCCGCGGCAACGTCAGTCTCAACIGGCGCAAGGAGGCGCCACATCGCCAG 1197
Db 1144 AAGGCCCCCGCGGCGATCGTCAAGTCTCAACIGGCGCAAGGAGGCGCCACATCGCCCGC 1204
QY 1198 AACTGCGCGCGCCCGCGGAAAGAGGCTGCTGGAAGTGGGCAAGGAGGCGGCGACAGATG 1257
Db 1204 AACTGCGCGCGCCCGCGGCAAGAGGCTGCTGGAAGTGGGCAAGGAGGCGGCGACAGATG 1263
QY 1258 AAGGACTGCACGAGGCGCGGAGGCAACTTCTGGGCAAGATCTGGGCGGCGGCGGCGGCG 1317
Db 1264 AAGGACTGCACGAGGCGCGGAGGCAACTTCTGGGCAAGATCTGGGCGGCGGCGGCGGCG 1323
QY 1318 CGCCCGCGCAACTTCTGCGAGAACCGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1377
Db 1324 CGCCCGCGCAACTTCTGCGAGGCGG-----CGCCGAGCGCGCGCGCGCGCGCGCGCG 1365
QY 1378 CCCCCGCGGAGAGCTTCGCTTCGAGGAGACCGCCCGCGCGCGCGCGCGCGCGCGCGCGCG 1437
Db 1366 CCCCCGCGGAGAGCTTCGCTTCGAGGAGACCGCCCGCGCGCGCGCGCGCGCGCGCGCGCG 1425
QY 1438 GACCGCGAGCCCTACCGCGAGCCCTGACCGCGCTGCGGAGCCTGTTGCGGAGCGCGCGCC 1497
Db 1426 G-----ACCGGAGACCCCTGACCGAGCCCTGAAGAGCCCTGTTGCGGCAACGACCCCG 1473
QY 1498 CTGAGCCAGTAA 1509
Db 1474 CTGAGCCAGAAA 1485
```

```
Sequence 16, Application US/10190435
Publication No. US20030143248A1
GENERAL INFORMATION:
APPLICANT: ZUR MEGEDE, Jan
APPLICANT: BARNETT, Susan W.
APPLICANT: LIAN, Ying
APPLICANT: ENGELBRECHT, Susan
APPLICANT: VAN RENSBURG, Estrelita J.
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
FILE REFERENCE: PPI8133.003 / 2302-18133
CURRENT APPLICATION NUMBER: US/10/190,435
CURRENT FILING DATE: 2002-12-30
NUMBER OF SEQ ID NOS: 319
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 16
LENGTH: 3462
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: GagProteinTrmL_C
US-10-190-435-16

Query Match      84.5%; Score 1275.2; DB 12; Length 3462;
Rest Local Similarity 92.5%; Pred. No. 5.4e-252;
Matches 1398; Conservative 0; Mismatches 78; Indels 36; Gaps 4;

QY 1 ATGGGGCGCGCGCCAGCATCTCTGGCGGCGAGAAAGCTGGACAAAGTGGGAGAAATCCGC 60
Db 7 ATGGGGCGCGCGCGCCAGCATCTCTGGCGGCGGCGGCAAGCTGGAGCGCTGGAGCGCATCCGC 66
QY 61 CTGGCGCGCGCGCGGCAAGAAAGCACTACATGCTGAAGCACTGGTGTGGCGACGCGCGAG 120
Db 67 CTGGCGCGCGCGCGGCAAGAAAGTGTACATGATGAAGCACTGGTGTGGCGACGCGCGAG 126
QY 121 CTGGAGGGCTTCGCCCTGAACCCCGCGGCTGCTGGAGACCGCGGAGGCTGCAAGCAGATC 180
Db 127 CTGGAGAAAGTTCGCCCTGAACCCCGCGGCTGCTGGAGACCGCGGAGGCTGCAAGCAGATC 186
QY 181 ATGAAGCAGCTGCAGCCCGCGCTGCAGACCGCGCACCGAGAGCTGCGCAGCCTGTACAAC 240
Db 187 ATCCGCCAGCTGCACCCCGCGCTGCAGACCGCGCACCGAGAGCTGAAGAGCCTGTTCAAC 246
QY 241 ACCGTGGCGCACCCCTGTACTGCTGCACCGCGCGCATCGAGTCCGCGACACCAAGGAGCC 300
Db 247 ACCGTGGCGCACCCCTGTACTGCTGCACCGAGAAAGATCGAGTCCGCGACACCAAGGAGCC 306
QY 301 CTGGACAAGATCGAGGAGGAGGAGCAACAAGTCCAGCAAGAACCCAGAGCCGCGGAGAG 360
Db 307 CTGGACAAGATCGAGGAGGAGGAGCAACAAGTCCAGCAAGAACCTCCAGAGAGATCCAGGCGGAGGCC 366
QY 361 GCGGAC--GGCAAGGTGAGCGAGCAAACTACCCCATCGTGCAGAAACCTGCGAGGCGCAGATG 417
Db 367 GCGGACAAGGCGCAAGGTGAGCGAGCAAACTACCCCATCGTGCAGAAACCTGCGAGGCGCAGATG 426
QY 418 GTGCACCGAGGCGCATCAGCCCGCGCACCCCTGAACCGCTGGTGAAGGTGATCGAGGAGAG 477
Db 427 GTGCACCGAGGCGCATCAGCCCGCGCACCCCTGAACCGCTGGTGAAGGTGATCGAGGAGAG 486
QY 478 GCGTTCAGCCCGGAGGTGATCCCGCATGTTCAACCGCGCTGAGCGAGGCGCGCCACCCCGAG 537
Db 487 GCGTTCAGCCCGGAGGTGATCCCGCATGTTCAACCGCGCTGAGCGAGGCGCGCCACCCCGAG 546
QY 538 GACCTGAACACAGATGTTGAACACCGTGGGCGGCGCCACCGAGCGCGCATGCAAGTGTGAAG 597
Db 547 GACCTGAACACAGATGTTGAACACCGTGGGCGGCGCCACCGAGCGCGCATGCAAGTGTGAAG 606
QY 598 GACACCATCAACGAGGAGGCGCGGAGTGGGACCGCGCTGCACCCCGTGCAGGCGCGGCCCG 657
Db 607 GACACCATCAACGAGGAGGCGCGGAGTGGGACCGCGCTGCACCCCGTGCAGGCGCGGCCCG 666
QY 658 GTGGCCCCCGCGGCGAGATGCGCGGACCCCGCGGCGGCGAGCATGCGCGCGCGCGCCACCGACCC 717
Db 658 GTGGCCCCCGCGGCGAGATGCGCGGACCCCGCGGCGGCGAGCATGCGCGCGCGCGCCACCGACCC 717
```


Db 847 CTGGACATCAAGCAGGGGCCCCAAGAGGCCCTTCGCGGACTAGCTGGACCGGCTCTTCAAG 936
QY 898 ACCGTGGCGGCGGAGCAGGCCCCAGGAGCGTGAAGAACTGATGACCGAGACCGCTGCTG 937
Db 907 ACCGTGGCGGCGGAGCAGGCCCCAGGAGGTTGAAGAACTGGATGACCGACACCGCTGCTG 938
QY 958 GTGCAGAAACGCCAACCCCCGACTGCAAGACCATCTCTGCGCGCTCTGGGCCCCCGGCGCCACC 1017
Db 967 GTGCAGAAACGCCAACCCCCGACTGCAAGACCATCTCTGCGCGCTCTGGGCCCCCGGCGCCACC 1026
QY 1018 CTGGAGGAGATGATGACCGGCTGCGCAGGGGCGGTGGGCGGCCCCCGCCACAAAGGCGCGGCTG 1077
Db 1027 CTGGAGGAGATGATGACCGGCTGCGCAGGGGCGGTGGGCGGCCCCCGCCACAAAGGCGCGGCTG 1086
QY 1078 CTGGCGGAGGGGATGAGCCAGGCCAACAGCGGTGAACATCATGATGACAGAGAGCAACTTC 1137
Db 1087 CTGGCGGAGGGGATGAGCCAGGCCAACACC--AGCGTGATGATGACAGAGAGCAACTTC 1143
QY 1138 AAGGGCCCCCGGCCAACGTCAAGTGTCTCAACTGGGGGCAAGGAGGGGCCACATCGGCCAAG 1197
Db 1144 AAGGGCCCCCGGCCGATCTCAAGTGTCTCAACTGGGGGCAAGGAGGGGCCACATCGGCCG 1203
QY 1198 AACTGCGCGGCGGCCCGCAAGAGGGGCTGTGTGAAGTGGGCAAGGAGGGGCCACCGAGATG 1257
Db 1204 AACTGCGCGGCGGCCCGCAAGAGGGGCTGTGTGAAGTGGGCAAGGAGGGGCCACCGAGATG 1263
QY 1258 AAGGACTGCACCGAGGCCAGGCCAACTTCTGTGGCAAGATCTGGGCCAGGCCAGCCAAAGGGC 1317
Db 1264 AAGGACTGCACCGAGGCCAGGCCAACTTCTGTGGCAAGATCTGGGCCAGGCCAGCCAAAGGGC 1323
QY 1318 CGCGCGCGCAACTTCTGTGCAGAACCGGACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1377
Db 1324 CGCGCGCGCAACTTCTGTGCAGAGCGG-----CGCGGAGCGCGCGCGCGCGCGCGCGCG 1385
QY 1378 CCGCGCGCGGAGAGCTTCGGCTTCGAGGAGACCGACCGCGCGCGCGCGCGCGCGCGCGCGCG 1437
Db 1366 CCGCGCGCGGAGAGCTTCGGCTTCGAGGAGACCGACCGCGCGCGCGCGCGCGCGCGCGCGCG 1425
QY 1438 GACCGGAGCGCCTACCGGAGCGCCCTGACCGCGCGCTGGCGAGCGCTGTGGCAAGCGCGCG 1497
Db 1426 G-----ACCGGAGACCGCTGACCGCGCGCTGAGAGAGCGCTGTGGCAAGCGCGCGCG 1475
QY 1498 CTGAGCCAGTAA 1509
Db 1474 CTGAGCCAGAAA 1485

RESULT 7

US-10-190-435-35
; Sequence 35, Application 9S/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEYER, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: P18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 35
; LENGTH: 4483
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: gp160mod.TV1.dV1V2-gagmod.BW965

US-10-190-435-35

Query Match 84.5%; Score 1275.2; DB 12; Length 4483;
Best Local Similarity 92.5%; Pred. No. 5.4e-252;
Matches 1398; Conservative 0; Mismatches 78; Indels 36; Gaps 4;
QY 1 ATGGGCG 60
Db 3005 ATGGGCG 3064
QY 61 CTGGCG 120
Db 3065 CTGGCG 3124
QY 121 CTGGAGGAGCTTGGCG 180
Db 3125 CTGGAGAGTTCG 3184
QY 181 ATGAAGCAGCTGCAGCG 240
Db 3185 ATCGCGCAGCTGCACCG 3244
QY 241 ACCGTGGCG 300
Db 3245 ACCGTGGCG 3304
QY 301 CTGGACAAGATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
Db 3305 CTGGACAAGATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3364
QY 361 GCGGAC---GGCAAGGTGAGCGCAGAACTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 417
Db 3365 GCGGACAAGGCAAGGTGAGCGCAGAACTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3424
QY 418 GTGCACCGAGCG 477
Db 3425 GTGCACCGAGCG 3484
QY 478 GCCTTCAGCG 537
Db 3485 GCCTTCAGCG 3544
QY 538 GACCTGAACACGAGTGTGAACACCGGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 597
Db 3545 GACCTGAACACGAGTGTGAACACCGGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3604
QY 598 GACACCAITACAGCAGGAGGCG 657
Db 3605 GACACCAITACAGCAGGAGGCG 3664
QY 656 GTGGCG 717
Db 3665 ATCG 3724
QY 718 CTGCAGGAGCAGATCGCGCTGGATGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 777
Db 3725 CTGCAGGAGCAGATCGCGCTGGATGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3784
QY 778 AACCGGTGGATCATCTCTGGCG 837
Db 3785 AACCGGTGGATCATCTCTGGCG 3844
QY 838 CTGCACATCCCG 897
Db 3845 CTGCACATCAACGAGGCG 3904
QY 898 ACCCTGCG 957
Db 3905 ACCCTGCG 3964
QY 958 GTGCAGAGACGCCAACCG 1017
Db 3965 GTGCAGAGACGCCAACCG 4024
QY 1018 CTGGAGGAGATGATGACCG 1077

Db 4025 CTGGAGGAGATGATGACCGCTGCCAGGGCGTGGGCGGCCCGCCAGCCAAAGGCCCGGIG 4084
QY 1078 CTGGCGGAGCGGATGAGCCAGGCCAACAGCGTGAACATCATGATGAGAGAGAACTTC 1137
Db 4085 CTGGCGGAGCGGATGAGCCAGGCCAACACC---AGCGTGAATGATGAGAGAGAACTTC 4141
QY 1138 AAGGGCCCGCGCGCAACGCTCAAGTGTCTCAACTGCGGCAAGGAGGCCACATCGCCCGC 1197
Db 4142 AAGGGCCCGCGCGCATGCTCAAGTGTCTCAACTGCGGCAAGGAGGCCACATCGCCCGC 4201
QY 1198 AACIGCGCGCGCCCGCGCAAGAGGCTGCTGGAAGTGGGCAAGGAGGGCCACAGATG 1257
Db 4202 AACTGCGCGCGCCCGCGCAAGAGGCTGCTGGAAGTGGGCAAGGAGGGCCACAGATG 4261
QY 1258 AAGGACTGCACCGAGCGCGCAAGGCTGCTGGAAGTGGGCAAGTCTGGCCAGCCACAGGCG 1317
Db 4262 AAGGACTGCACCGAGCGCGCAAGGCTGCTGGAAGTGGGCAAGTCTGGCCAGCCACAGGCG 4321
QY 1318 CGCCCGCGCAACTTCTGTCAGAACCGCAGCGAGCGCGCGCGCGCGCGCGCGCGCGCG 1377
Db 4322 CGCCCGCGCAACTTCTGTCAGAGCGG-----CGCCGAGCGCGCGCGCGCGCGCGCG 4353
QY 1378 CCGCGCGCGGAGACTTCGGCTTCGAGAGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1437
Db 4364 CCGCGCGCGGAGAGCTTCGGCTTCGAGAGAGACCGCGCGCGCGCGCGCGCGCGCGCGCG 4423
QY 1438 GACCGGAGCGCCCTACCGCGAGCGCCCTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1497
Db 4424 G-----ACCGCGAGACCCCTGACCGAGCGCTGGAAGAGCCCTGTCGGCAAGCGCGCG 4471
QY 1498 CTGAGCGAGTAA 1509
Db 4472 CTGAGCGCAATAA 4483

RESULT 8
US-10-190-435-34
; Sequence 34, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEYER, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190.435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34
; LENGTH: 4606
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: gp160mod.1Vi.dv1.gagmod.HW965
US-10-190-435-34

Query Match 84.5%; Score 1275.2; DB 12; Length 4606;
Best Local Similarity 92.5%; Pred. No. 5.4e-252;
Matches 1398; Conservative 0; Mismatches 78; Indels 36; Gaps 4;
QY 1 ATGGCGCGCGCGCGCGCATCTCGCGCGCGGAGAGCTGGACAGTGGGAGAGATCCCG 60
Db 3128 ATGGCGCGCGCGCGCGCATCTCGCGCGCGGAGAGCTGGACCGCTGGGAGCGCATCCCG 3187
QY 61 CTGGCGCGCGCGCGGCAAGAGCACTACATGCTGAAGCACCTGGTGGGCGCGCGCGAG 120
Db 3188 CTGGCGCGCGCGCGGCAAGAGTGTCTACATGATGAAGCACCTGTGTGGGCGCGCGCGAG 3247

QY 121 CTGGAGGGCTTCGCGCCCTGAACCCCGCGCTGCTGGAGACCGCGGAGGGCTGCAAGCAGATC 180
Db 3248 CTGGAGAAAGTTCCCGCTGAACCCCGCGCTGCTGGAGACCGAGGAGGGCTGCAAGCAGATC 3307
QY 181 ATGAAGCAGCTGCAGCCCGCTGCAGACCGGCACCGAGAGCTGCGAGCTGTACAAC 240
Db 3308 ATCCCGCAGCTGCACCCCGCTGCAGACCGGCAGCGAGAGCTGAAGAGCTGTTCAAC 3367
QY 241 ACCTGTGCCACCTGTACTGCTGCACCGCGCATCGAGGTCCCGGACACACCAAGGAGGCC 300
Db 3368 ACCGTGGCCACCTGTACTGCTGCACGAGAGATCGAGGTCCCGGACACCAAGGAGGCC 3427
QY 301 CTGGACNAGATCGAGGAGGAGCAACAAGTCCACGAGAGACCCAGCAGGCCCAAGGAG 360
Db 3428 CTGGACNAGATCGAGGAGGAGCAACAAGTCCACGAGAGATCGAGGTCCCGGACCGAGGCC 3487
QY 361 GCGGAC---GGCAAGGTGAGCCAGAACTACCGCATCGTGCAGAACTTCAGCGCCAGATG 417
Db 3488 GCGGACAAAGGCAAGGTGAGCCAGAACTACCGCATCGTGCAGAACTTCAGCGCCAGATG 3547
QY 418 CTGCACCGGCGCATCAGCCCGCGCACCTGAAGCGCTGGTGAAGGTGATCGAGGAGAAG 477
Db 3548 GTGCACCGGCGCATCAGCCCGCGCACCTGAAGCGCTGGTGAAGGTGATCGAGGAGAAG 3607
QY 478 GCTTCAGCCCGGAGGTGATCCCGCATGTTACCGCGCTTGAAGGAGGGCGCCACCCCGCAG 537
Db 3608 GCTTCAGCCCGGAGGTGATCCCGCATGTTACCGCGCTTGAAGGAGGGCGCCACCCCGCAG 3667
QY 538 GACCTGAACACCAAGTGAACACCGTGGCGGCCACCAAGGCGCCATGACAGATGCTGAAG 597
Db 3668 GACCTGAACACCAAGTGAACACCGTGGCGGCCACCAAGGCGCCATGACAGATGCTGAAG 3727
QY 598 GACACCATCAACGAGGAGGCGCGCGAGTGGGAGCGCGCTGCACCCCGTGCAGCGCGCGCCC 657
Db 3728 GACACCATCAACGAGGAGGCGCGCGAGTGGGAGCGCGCTGCACCCCGTGCAGCGCGCGCCC 3787
QY 658 GTGGCCCCCGCGCAGATGCGCGACCCCGCGCGGAGCGAGATCGCGCGCGCGCGCGCGCGC 717
Db 3788 ATCGCCCCCGCGCAGATGCGCGAGCCCGCGCGGAGCGAGATCGCGCGCGCGCGCGCGCGC 3847
QY 718 CTGCAGGAGCAGATCGCTGGATGACCAAGACCCCGCGCGTGCCTGCGCGCGCGCGCGCGC 777
Db 3848 CTGCAGGAGCAGATCGCTGGATGACCAAGACCCCGCGCGTGCCTGCGCGCGCGCGCGCGC 3907
QY 778 AAGCGGTGGATCATCTGGCGCTGAACAAGATCGTGGGATGTATACAGCCCGCTGAGCAIC 837
Db 3908 AAGCGGTGGATCATCTGGCGCTGAACAAGATCGTGGGATGTATACAGCCCGCTGAGCATC 3967
QY 838 CTGGACATCCGCGCAGGCGCCCAAGGAGCGCTTCGCGGAC*ACGTGGACCGCTTCTTCAAG 897
Db 3968 CTGGACATCAAGCAGGCGCCCAAGGAGCGCTTCGCGGAC*ACGTGGACCGCTTCTTCAAG 4027
QY 898 ACCCTGCGCGCGAGCAGCGCCACCGAGCGCGCGCGTGAAGAACTGGATGACCGACCGCTGCTG 957
Db 4028 ACCCTGCGCGCGAGCAGCAGCGCCACCGAGCGCGCGTGAAGAACTGGATGACCGACCGCTGCTG 4087
QY 958 GTGCAGAACCGCAACCCCGACTGCAAGACCATCTCGCGCGCTTCGCGCGCGCGCGCGCGCGC 1017
Db 4088 GTGCAGAACCGCAACCCCGACTGCAAGACCATCTCGCGCGCTTCGCGCGCGCGCGCGCGCGC 4147
QY 1018 CTGGAGGAGATGATGACCGCTGCGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1077
Db 4148 CTGGAGGAGATGATGACCGCTGCGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4207
QY 1078 CTGGCGGAGGCGATGAGCCAGGCCAACAGCGTGAACATCATGATGAGAGAGCAACTTC 1137
Db 4208 CTGGCGGAGGCGATGAGCCAGGCCAACACC---AGCGTGA*GATGAGAGAGCAACTTC 4264
QY 1138 AAGGGCCCCCGCGCAACGTCAAGTGTCTCAACTGCGGCAAGGAGGGCGCGCGCGCGCGCG 1197
Db 4265 AAGGGCCCCCGCGCATCGTCAAGTGTCTCAACTGCGGCAAGGAGGGCGCGCGCGCGCGCG 4324

QY 1378 CCGCCGCGGAGAGCTTCGGCTCGAGGAGACACCCCGCCCAAGCAGGAGCCCAAG 1437
|||||
Db 4496 CCGCCGCGGAGAGCTTCGGCTCGAGGAGACACCCCGCCCAAGCAGGAGCAAG 4555
|||||
QY 1438 GACCGGAGCCCTACCGGAGCCCTGACCGCCCTGCGCAGCCTGTTCCGAGCGGCC 1497
|||||
Db 4556 G-----ACCGGAGACCCTGACCAGCCTGAAGAGCCCTGTTCCGCAACGACCC 4603
|||||
QY 1498 CTGAGCCAGTAA 1509
|||||
Db 4604 CTGAGCCAAATA 4615
|||||
RESULT 10
US-10-190-435-38
; Sequence 38, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZJR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBERG, Estrelita J.
; TITLE OF INVENTION: POLYPEPTIDES ENCODING ANTIGENIC HIV TYPE 1
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: P18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; NUMBER OF SEQ ID NOS: 319
; NUMBER OF FILING DATE: 2002-12-30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38
; LENGTH: 4702
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: gp160mod.TV1-gagmod.BW955
US-10-190-435-38
Query Match 84.5%; Score 1275.2; DB 12; Length 4702;
Best Local Similarity 92.5%; Prob. No. 5.4e-252;
Matches 1398; Conservative 0; Mismatches 75; Indels 36; Gaps 4;
QY 1 ATGGGCGCGCGCCAGCATCTTCGGCGCGGCGGAGAAAGCTGGACAAGTGGGAGAGATCGGC 60
|||||
Db 3224 ATGGGCGCGCGCCAGCATCTTCGGCGCGGCGGCAAGCTGGACCCCTGGGAGCGCATCGGC 3283
|||||
QY 61 CTGCGCCCGCGCGGCAAGACACTACATGCTGAAGCACCTGTGTGGGCCAGCCGCGAG 120
|||||
Db 3284 CTGCGCCCGCGCGGCAAGAGTGTACATGATGAAGCACCTGTGTGGGCCAGCCGCGAG 3343
|||||
QY 121 CTGAGGGCTTCGCCCTGAACCCCGGCTGTGTGAGACCCGCGAGGGCTGCAAGCAGATC 180
|||||
Db 3344 CTGAGAGAGTTCGCCCTGAACCCCGGCTGTGTGAGACCCAGCGAGGGCTGCAAGCAGATC 3403
|||||
QY 181 ATGAAGCAGCTGCAGCCCGCCCTGCAGACCGGACCGGAGGAGTGGCAGCCCTGACAAAC 240
|||||
Db 3404 ATCGGCCAGCTGCACCCCGCCCTGCAGACCGGCGAGCGAGGAGTGAAGAGCCCTGTTCAAC 3463
|||||
QY 241 ACCGTGGCCACCTGTACTGCGTGCACGCGCGGATCGAGGTCCGCGACACCAAGGAGGCC 320
|||||
Db 3464 ACCGTGGCCACCTGTACTGCGTGCACGAGAGATCGAGGTCCGCGACACCAAGGAGGCC 3523
|||||
QY 301 CTGACAAAGATCGAGGAGGAGGACAAACAAGTCCAGCAGAAAGACCCAGCGCCCAAGGAG 360
|||||
Db 3524 CTGACAAAGATCGAGGAGGAGGACAAACAAGTCCAGCAGAAAGTCCAGCAGCGCGAGGCC 3583
|||||
QY 361 GCGGAC---GGCAGGTCAGCCCAAGAACTACCCCATCGTGCAGAACCTGCAGGCGCCAGAAG 417
|||||
Db 3584 GCGGACAAAGGTCAGGTCAGGACAACTACCCCATCGTGCAGAACCTGCAGGCGCCAGAAG 4643
|||||
QY 418 GTGCACAGGCCCATCAGCCCCCGCCACCCCTGAACCCCTGGGTGAAGGTGATCGAGGAGAA 477
|||||

Db 3644 GTGCACAGGCCCAICAGCCCCCGCACCCCTGAACCCCTGGGTGAAGGTGATCGAGGAGAA 3703
|||||
QY 478 GCCTTACGCCCGGAGGTGATCCCAATGTTACCGCCCTGAGCGAGGCGCCACACCCCGCAG 537
|||||
Db 3704 GCCTTACGCCCGGAGGTGATCCCAATGTTACCGCCCTGAGCGAGGCGCCACCCCGCAG 3763
|||||
QY 538 GACCTGAACACGATGTTGAACACCGTGGCGGCGCACAGGCGCCCATGCAAGATGCTGAAG 597
|||||
Db 3764 GACCTGAACACGATGTTGAACACCGTGGCGGCGCACAGGCGCCCATGCAAGATGCTGAAG 3823
|||||
QY 598 GACACCATCAACGAGGAGGCGCGCGAGTGGGACCGCTGCACCCCTGCAGCGCCCGGCC 657
|||||
Db 3824 GACACCATCAACGAGGAGGCGCGCGAGTGGGACCGCTGCACCCCTGCAGCGCCCGGCC 3883
|||||
QY 658 GTGGCCCCCGGCCAGATGCCCGACCCCGCGCGGAGGAGACATCGCCGGCGCCACCGACCC 717
|||||
Db 3884 ATCGCCCCCGGCCAGATGCCCGAGCCCGCGCGGAGGAGACATCGCCGGCGCCACCGACCC 3943
|||||
QY 718 CTGCAGGAGCAGATCGCCTGGATGACCAAGACACCCCGCGTGGCGGCGCCACATCTAC 777
|||||
Db 3944 CTGCAGGAGCAGATCGCCTGGATGACCAAGACACCCCGCGTGGCGGCGCCACATCTAC 4003
|||||
QY 778 AAGCGGTGGATCATCTTGGGCTGAACAAGATCTGCGGATGTACAGCCCGCGTGAATC 837
|||||
Db 4004 AAGCGGTGGATCATCTTGGGCTGAACAAGATCTGCGGATGTACAGCCCGCGTGAATC 4063
|||||
QY 838 CTGGACATCCGCCCGGAGGCGCCCAAGGAGCCCTTCGCGGACTACGTGCGGCTCTTCAAG 897
|||||
Db 4064 CTGGACATCAAGCAGGCGCCCAAGGAGCCCTTCGCGGACTACGTGCGGCTCTTCAAG 4123
|||||
QY 898 ACCCTGGCGCGCGGAGCGCCACCCAGGAGCGTGAAGAACTGGATGACCGAGACCCCTGCTG 957
|||||
Db 4124 ACCCTGGCGCGCGGAGCGCCACCCAGGAGCGTGAAGAACTGGATGACCGAGACCCCTGCTG 4183
|||||
QY 958 GTGCAGAACGCCCAACCCCGACTGCAACACCATCTCTGCGGCTCTCGGCCCGCGCGCACCC 1017
|||||
Db 4184 GTGCAGAACGCCCAACCCCGACTGCAACACCATCTCTGCGGCTCTCGGCCCGCGCGCACCC 4243
|||||
QY 1018 CTGGAGGAGATGATGACCGCCCTGCCAGGGCGTGGCGGCGCCCGGCGCAAGCGCCGCTG 1077
|||||
Db 4244 CTGGAGGAGATGATGACCGCCCTGCCAGGGCGTGGCGGCGCCCGGCGCAAGCGCCGCTG 4303
|||||
QY 1078 CTGGCCGAGCGGATGAGCCAGGCCCAACAGCGCTGAACATCATATGATGCAAGAGCAACTTC 1137
|||||
Db 4304 CTGGCCGAGCGGATGAGCCAGGCCCAACACCC---ACCGTATGATGCAAGAGAGCAACTTC 4360
|||||
QY 1138 AAGGGCCCCCGCGGCAACGTCAAGTCTCAACTGCGGCAAGGAGGGCGCCACATCGCCCAAG 1197
|||||
Db 4361 AAGGGCCCCCGCGGCAATCTCAAGTCTCAACTGCGGCAAGGAGGGCGCCACATCGCCCGC 4420
|||||
QY 1198 AACTGCGCGCGCCCGCGCAAGAGGGCTGCTGGAAGTGGGCAAGGAGGGCGCCACAGATG 1257
|||||
Db 4421 AACTGCGCGCGCCCGCGCAAGAGGGCTGCTGGAAGTGGGCAAGGAGGGCGCCACAGATG 4480
|||||
QY 1258 AAGGACTGCACCGAGCGCCAGGCCCAACTCTCTGGGCAAGATCTGGCCAGCCACAGGGCC 1317
|||||
Db 4481 AAGGACTGCACCGAGCGCGCCAGGCCCAACTCTCTGGGCAAGATCTGGCCAGCCACAGGGCC 4540
|||||
QY 1318 CGCCCCGGCAACTCTCTGAGAACCGGAGCGCGCGCGCCCGCCCGCCCGTGGCCCGCCCGCC 1377
|||||
Db 4541 CGCCCCGGCAACTCTCTGAGAGCGCG-----CCCCGAGCCCCACCGCC 4582
|||||
QY 1378 CCCCCCGCGGAGAGCTTCGCTTCGAGGAGACCAACCCCGCCCGCCCGGAGGAGCGCCCAAG 1437
|||||
Db 4583 CCCCCCGCGGAGAGCTTCGCTTCGAGGAGACCAACCCCGCGCGCCAGAGGAGGAGGAGCAAG 4642
|||||
QY 1438 GACCCGAGCCCTACCGCGAGCGCCCTGACCGCCCTGCGGAGCCCTGTCGCGAGCGGCGCC 1497
|||||
Db 4643 G-----ACCGGAGAGCCCTGACAGCCCTGAAGAGCCCTGTCGCGCAACGACCGCC 4690
|||||
QY 1498 CTGAGCCAGTAA 1509
|||||

Db 4692 CTGAGCCCATAA 4702

RESULT 11

US-10-190-435-17
; Sequence 17, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:

; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PPI8133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 17
; LENGTH: 4716

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence:

; OTHER INFORMATION: GagProtInaRTmuttrNef_C

US-10-190-435-17

Query Match 84.5%; Score 1275.2; DB 12; Length 4716;

Best Local Similarity 92.5%; Pred. No. 5.4e-252;

Matches 1398; Conservative 0; Mismatches 78; Indels 36; Gaps 4;

QY 1 ATGGGGCGCGCGCCAGCATCCCTGGCGGCGAGAAAGCTGGACAAAGTGGAGAGATCCGC 60
Db 7 ATGGCGCGCGCGCGCGCATCCCTGGCGGCGGCGAGCTGGACGCTGGAGCGCATCCGC 66
QY 61 CTGCGCGCGCGCGCGCAAGAAAGCACTACATGCTGAAGCACTGGTGTGGCGCAGCGCGAG 120
Db 67 CTGCGCGCGCGCGCGCAAGAAAGTGTACATGATGAAGCACTGGTGTGGCGCAGCGCGAG 126
QY 121 CTGGAGGGGCTCGCCCTGAACCCCGCGCTGCTGGAGAGCGCGCGAGGGGTGCAAGCAGATC 180
Db 127 CTGGAGAAATCGCCCTGAACCCCGCGCTGCTGGAGAGCGCGAGGGGTGCAAGCAGATC 186
QY 181 ATGAAGCACTGCAGCGCGCGCGCTGCAGACCGCGCACCGAGAGAGCTGCCAGCGCTGTACAA 240
Db 187 ATCGCGCACTGCACCGCGCGCGCTGCAGACCGCGCGCAGCGAGAGCTGAAGAGCGCTGTAA 246
QY 241 ACCGTGGCACCCCTGTACTGGGTGCACCGCGCGCATTCAGGTCCCGGACACCAAGAGGCC 300
Db 247 ACCGTGGCACCCCTGTACTGGGTGCACCGAGAAAGATCGAGGTCCCGGACACCAAGAGGCC 306
QY 301 CTGGACAAAGATCGAGGAGGAGCAGAAAGTCCCGAGCAAGACCCAGCAGCGCGCAAGAG 360
Db 307 CTGGACAAAGATCGAGGAGGAGCAGCAACAAGTCCCGAGCAAGATCCAGAGCGCGAGGCC 366
QY 361 GCCGAC---GGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACTTCAGGGCGCAGATG 417
Db 367 GCCGACAAAGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACTTCAGGGCGCAGATG 426
QY 418 GTGCACCAAGCCATCAGCCCGCGCACCCCTGAACGCTTGGGTGAAGGTGATCGAGGAGAG 477
Db 427 GTGCACCAAGCCATCAGCCCGCGCACCCCTGAACGCTTGGGTGAAGGTGATCGAGGAGAG 486
QY 478 GCCTTCAGCCCGGAGGTGATCCCCCATGTTACCCGCCCTGAGCGAGGGCGCCACCCCGCAG 537
Db 487 GCCTTCAGCCCGGAGGTGATCCCCCATGTTACCCGCCCTGAGCGAGGGCGCCACCCCGCAG 546
QY 538 GACCTGAACACGATGTTGAACACCGTGGGGCGGCCACCGCGCCCATGAGATGCTGAAG 597
Db 547 GACCTGAACACGATGTTGAACACCGTGGGGCGGCCACCGCGCCCATGAGATGCTGAAG 606

QY 598 GACACCATCAACGAGAGGAGCGCGCGAGTGGAGCGCGCTGCAGCCCGCTGCAGCGCGCGCC 657
Db 607 GACACCATCAACGAGAGGAGCGCGCGAGTGGAGCGCGCTGCAGCCCGCTGCAGCGCGCGCC 666
QY 658 GTGGCGCGCGCGCAGATCGCGACCCCGCGCGGCGAGCGACATCGCGCGCGCGCGCAGCACC 717
Db 667 ATCGCGCGCGCGCAGATCGCGACCCCGCGCGGCGAGCGACATCGCGCGCGCGCAGCACC 726
QY 718 CTGAGGAGCAGATCGCTGGATGACCAAGCAACCCCGCGCTGCGCGCGTGGGCGACATCTAC 777
Db 727 CTGAGGAGCAGATCGCTGGATGACCAAGCAACCCCGCGCTGCGCGCGTGGGCGACATCTAC 786
QY 778 AAGCGGTGGATCATCTGGGCTGAACAAGATCGTGGGATGACAGCCCGCTGAGCATC 837
Db 787 AAGCGGTGGATCATCTGGGCTGAACAAGATCGTGGGATGACAGCCCGCTGAGCATC 846
QY 838 CTGGACATCCGCGAGGCGCGCGCGCGCGCGCGCTGCGGAGTACGTGGACCGCTTCTCAAG 897
Db 847 CTGGACATCAACGAGGCGCGCGCGCGCGCGCGCTTCCGCGACTACGTGGACCGCTTCTCAAG 906
QY 898 ACCCTGCGCGCGCGAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 957
Db 907 ACCCTGCGCGCGCGAGCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 966
QY 958 GTGAGAAAGCG 1017
Db 967 GTGAGAAAGCG 1026
QY 1018 CTGGAGGAGATGATGACCGCGCTGCCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1077
Db 1027 CTGGAGGAGATGATGACCGCGCTGCCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1086
QY 1078 CTGGCGGAGGCGATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1137
Db 1087 CTGGCGGAGGCGATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1143
QY 1138 AAGGGCG 1197
Db 1144 AAGGGCG 1203
QY 1198 AACTGCG 1257
Db 1204 AACTGCG 1263
QY 1258 AAGGACTGCACCGAGCG 1317
Db 1264 AAGGACTGCACCGAGCG 1323
QY 1318 CGSCCGCGCGCAACTCTCTGCAGAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1377
Db 1324 CGSCCGCGCGCAACTCTCTGCAGAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1365
QY 1378 CCCCCCGCGCGAGAGCTTCGGCTTCGAGGAGAGCAACCCCGCGCGCGCGCGCGCGCGCG 1437
Db 1366 CCCCCCGCGCGAGAGCTTCGGCTTCGAGGAGAGCAACCCCGCGCGCGCGCGCGCGCGCG 1425
QY 1438 GACCGCGAGCGCTACCG 1497
Db 1426 G-----ACCGCGAGAGCGCTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1473
QY 1498 CTGAGCCAGTAA 1509
Db 1474 CTGAGCCAGAAA 1485

RESULT 12

US-10-190-435-20
; Sequence 20, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying

```

; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PPI18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190.435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 2742
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GagTatRevNef_c
US-10-190-435-20

Query Match      84.4%; Score 1273.6; DB 12; Length 2742;
Best Local Similarity 92.4%; Pred. No. 1.2e-251;
Matches 1397; Conservative 0; Mismatches 79; Indels 36; Gaps 4;

QY      1 ATGGGCGCCGCGCCAGCATCTCTGGCGGCGAGAGCTGGACAAAGTGGGAGAGATCCGC 60
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      61 CTGCGCCCGCGCGGCAAGAACACTACATGCTGAAGCACTGTGTGGGCCAGCGCGAG 120
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      67 CTGCGCCCGCGCGGCAAGAGTGTACATGATGAAGCACTGTGTGGGCCAGCGCGAG 126
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      121 CTGGAGGATTCGCGCCCTGAACCCCGCGCTGCTGGAGACCGCGGAGGCTGCAGAGAGAT 180
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      127 CTGGAGAGATTCGCGCCCTGAACCCCGCGCTGCTGGAGACCGGAGGCTGCAAGCAATC 186
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      181 ATGAAGCAGTGCAGCCCGCGCTGCAGACCGGCACCGAGAGCTGCGCAGCCCTGTACAA 240
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      187 ATCCGCCAGTGCACCCCGCGCTGCAGACCGGCACCGAGAGCTGAAGAGCCCTGTCAAC 246
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      241 ACCGTGGCCACCCCTGTACTGCTGTCACCGCGGCATCGAGTCCGGCAGACCAAGAGGCC 300
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      247 ACCGTGGCCACCCCTGTACTGCTGTCACCGAGAGATCGAGGTCCGGCAGACCAAGAGGCC 306
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      301 CTGGACAAGATCGAGGAGGAGCAGACAAGTCCCAGCAGAACCCAGCAGGCCCAAGCAG 360
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      307 CTGGACAAGATCGAGGAGGAGCAGACAAGTCCACGACAGAGATCCAGCGCCCGAGGCC 366
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      361 GCGGAC---GCCAAGGTGAGCCAGAACTACCCCATCTGTGAGAACTGTGAGGCGCCAGAG 417
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      367 GCGGACAAGGCAAGGTGAGCCAGAACTACCCCATCTGTGAGAACTGTGAGGCGCCAGAG 426
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      418 GTGCACCAAGGCATCAGCCCGCGCATCTGAAGCCCTGGTGAAGGTGATCGAGGAGAGAG 477
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      427 GTGCACCAAGGCATCAGCCCGCGCATCTGAAGCCCTGGTGAAGGTGATCGAGGAGAGAG 486
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      478 GCCCTCAGCCCGGAGGTGATCCCATGTTTACCGCCCTGTAGCGGAGGGCGCCACCCCGCCAG 537
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      487 GCCCTCAGCCCGGAGGTGATCCCATGTTTACCGCCCTGTAGCGGAGGGCGCCACCCCGCCAG 546
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      538 GACCTGAACACGATGTTGAACACCGTGGGCGGCCACCGCCCATGCGAGATGCTGAAG 597
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      547 GACCTGAACACGATGTTGAACACCGTGGGCGGCCACCGCCCATGCGAGATGCTGAAG 606
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      598 GACACCATCAAGGAGGAGCGCGCGAGTGGACCGCTGTGACCCCGTGTAGCGCCCGCGCC 657
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      607 GACACCATCAAGGAGGAGCGCGCGAGTGGGACCGCTGTGACCCCGTGTGACCGCCCGCC 666
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      658 GTGGCCCCCGGCAGATGCGGCAACCCCGCGGACGAGCATGCGCGCGCGCCAGCAGCACC 717
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      667 ATCGCCCCCGGCAGATGCGGCAACCCCGCGGACGAGCATGCGCGCGCCACCGCAGCACC 726
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      718 CTGCAGGAGCAGATCGCTGTGATGACCAACCCCGCGTGGCCGTGGCGGACATCTAC 777
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      727 CTGCAGGAGCAGATCGCTGTGATGACCAACCCCGCGTGGCGGACATCTAC 786
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```

QY      778 AAGCGGTGGATCATCTCTGGGCTGAACAAGATCGTGGGATGTACAGCCCCGGTGGAGCATC 837
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      787 AAGCGGTGGATCATCTCTGGGCTGAACAAGATCGTGGGATGTACAGCCCCGGTGGAGCATC 846
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      838 CTGGACATCCGCGCAGGCCCCCAAGGAGCCCTCCGCGACTAGCTGGAGCCGCTTCTCAAG 897
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      847 CTGGACATCAAGCAGGCCCCCAAGGAGCCCTCCGCGACTAGCTGGAGCCGCTTCTCAAG 906
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      898 ACCCTGCGCGCGCAGCAGGCCACCCAGGAGCGTGAAGAACTGGAIGACCGGAGACCCTGCTG 957
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      907 ACCCTGCGCGCGCAGCAGGCCACCCAGGAGGTGAAGAACTGGAIGACCGGAGACCCTGCTG 966
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      958 GTGCAGAACCGCCACCCCGACTGCAAGACCACTCTCGCGCTCTCGGCCCGCGGCCACC 1017
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      967 GTGCAGAACCGCCACCCCGACTGCAAGACCACTCTCGCGCTCTCGGCCCGCGGCCACC 1026
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      1018 CTGGAGGAGATGATGACCGCTCCAGGGCGTGGGCGGCCCGGCCACCAAGGCCCGGCTG 1077
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      1027 CTGGAGGAGATGATGACCGCTCCAGGGCGTGGGCGGCCCGGCCACCAAGGCCCGGCTG 1086
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      1078 CTGGCCGAGGGGATGAGCCAGGCCAACACAGCGTGAACATCATGATGACAGAGCAACTTC 1137
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      1087 CTGGCCGAGGGGATGAGCCAGGCCAACACC---AGCGTGTGATGATGACAGAGCAACTTC 1143
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      1138 AAGGCCCCCGCGCGCAACGTCAAGTCTTCAACTGCGGCAAGAGGGGCCACATCGCCCAAG 1197
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      1144 AAGGCCCCCGCGCGCATCGTCAAGTCTTCAACTGCGGCAAGAGGGGCCACATCGCCCGC 1203
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      1198 AACTGCGCGCGCGCGCAAGRAGGCGTCTGCAAGTGGGCAAGGAGGGGCCACCATGATG 1257
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      1204 AACTGCGCGCGCGCGCAAGRAGGCGTCTGCAAGTGGGCAAGGAGGGGCCACCATGATG 1263
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      1258 AAGGACTGCACCGGAGCGCCAGGCCAACTTCTCTGGCAAGATCTGGGCCAGGCCACAAGGGC 1317
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      1264 AAGGACTGCACCGGAGCGCCAGGCCAACTTCTCTGGCAAGATCTGGGCCAGGCCACAAGGGC 1323
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      1318 CGCCCGCGCAACTTCTCTGCAGAACCGCAGCGAGCGCGCGCCCGCCACCGTSCCACCAGCC 1377
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      1324 CGCCCGCGCAACTTCTCTGCAGAGCGG-----CCCCGAGGCCACCGCC 1365
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      1378 CCCC CGCGGAGAGCTTCGGCTTCAGGAGAGCACACCCCGCCCGCCCAAGCAGGAGGCCCAAG 1437
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      1366 CCCC CGCGGAGAGCTTCGGCTTCAGGAGAGCACACCCCGCCCGCCCAAGCAGGAGAGCAAG 1425
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      1438 GACCGGAGCCCTACCGCGAGCCCTGACCGCCCTGCGCAGCCCTGTCGGCAGCGGCCCC 1497
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      1426 G-----ACCGCGAGACCCCTGACCAGCCTGAAGAGCGCTCTCGGCAACGACCCCC 1473
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      1498 CTGAGCCAGTAA 1509
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      1474 CTGAGCCCAAGAA 1485
DB      1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 13
US-10-190-435-9
; Sequence 9, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PPI18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190.435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 3930
; TYPE: DNA
```



```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial: Sequence: GagCompPolmutatt_C
US-10-190-435-9

Query Match      84.4%; Score 1273.6; DB 12; Length 3930;
Best Local Similarity 92.4%; Pred. No. 1.le-251;
Matches 1397; Conservative 0; Mismatches 79; Indels 36; Gaps 4;

QY 1 ATGGCGCCCGCCGCGCAAGACACTGCTCGCGCGGCGGAGAAAGCTGGACAACTGGGAGAGAAAGCC 60
Db 7 ATGGCGCCCGCCGCGCGCAAGACACTGCTCGCGCGGCGGAGAAAGCTGGGAGAGAAAGCCATCCGC 66
QY 61 CTGGCGCCCGCGCGGCAAGACACTACATGCTGAAGCACTGGTGTGGGCGGCGGAG 120
Db 57 CTGGCGCCCGCGCGGCAAGAGTGTACATGATGAAGCACTGGTGTGGGCGGCGGAG 126
QY 121 CTGGAGGGCTTCGCCCTGAACCCCGCCCTGCTGGAGACCGCGGAGGCTGCGAGCAGATC 180
Db 127 CTGGAGAAATTCCGCCCTGNAACCCCGCCCTGCTGGAGACCGCGGAGGCTGCGAGCAGATC 186
QY 181 ATGAAGCAGCTGACGCCCGCCCTGCAGACCGGCGGAGGAGCTGCCAGCCCTGTACAA 240
Db 187 ATCCGCCAGCTGACACCCCGCCCTGCAGACCGGCGGAGGAGTGAAGAGCTGTTCAC 246
QY 241 ACCGTGGCCACCTGTACTGCTGCTACCGCGGCGGATCGAGGTCCGCGACACCAAGAGGCC 300
Db 247 ACCGTGGCCACCTGTACTGCTGCTACCGCGGCGGATCGAGGTCCGCGACACCAAGAGGCC 306
QY 301 CTGGACAAGATCGAGCAGGAGCGAACAAGTCCAGCAGAAAGTCCAGCCCGCGAGCGCAAGGAG 360
Db 307 CTGGACAAGATCGAGCAGGAGCGAACAAGTCCAGCAGAAAGTCCAGCAGGCGGCGGAGGCC 366
QY 361 GCCAC---GGCAAGGAGCGCAAGAACTACCCCATCGTGCAGAACTCGCAGGCGGAGATC 417
Db 367 GCCACAAGGCGAAGGTGAGCCAGAACTACCCCATCGTGCAGAACTCGCAGGCGGAGATC 423
QY 418 GTGCACACGCGCCATCAGCCCGCGGCGACCCCTGAACCGCTGGTGAAGGTGATCGAGGAGAG 477
Db 427 GTGCACACGCGCCATCAGCCCGCGGCGACCCCTGAACCGCTGGTGAAGGTGATCGAGGAGAG 484
QY 478 GCTTCAGCCCGCGAGGTGATCCCATGTTTACCGCGCTGAGCGAGGCGCGCACCGCCCGCAG 537
Db 487 GCTTCAGCCCGCGAGGTGATCCCATGTTTACCGCGCTGAGCGAGGCGCGCACCGCCCGCAG 543
QY 538 GACCTGAACACGATGTGAACACGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 597
Db 547 GACCTGAACACGATGTGAACACGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 603
QY 598 GACACCATCAACGAGGAGCGCGGAGTGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 657
Db 607 GACACCATCAACGAGGAGCGCGGAGTGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 663
QY 658 GTGGCCCGCGCGGCGGAGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 717
Db 667 ATCGCCCGCGCGGCGGAGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 723
QY 718 CTGCAGGAGCAGATCGGCTGGATGACCAAGCAACCCCGCGGCGGCGGCGGCGGCGGCGGCG 777
Db 727 CTGCAGGAGCAGATCGGCTGGATGACCAAGCAACCCCGCGGCGGCGGCGGCGGCGGCGGCG 783
QY 778 AAGCGGTGGATCATCTGGGCGCTGAACAAGATCGTGGCGGATGTACAGCCCGCGTGAAG 837
Db 787 AAGCGGTGGATCATCTGGGCGCTGAACAAGATCGTGGCGGATGTACAGCCCGCGTGAAG 843
QY 838 CTGGACATCCGCCAGGCGCCCAAGGAGCCCTTCCGGGACTACGTGGACCGCTTCTCAAG 897
Db 847 CTGGACATCAAGCAGGCGCCCAAGGAGCCCTTCCGGGACTACGTGGACCGCTTCTCAAG 903
QY 898 ACCCTGGCGCGCGGAGGAGCGCCACCCAGGACGCTGAAGAACTGGATGACCGGAGCCCTGCTG 957
Db 907 ACCCTGGCGCGCGGAGGAGGAGCGCCACCCAGGAGGTGAAGAACTGGATGACCGGAGCCCTGCTG 963
```

```

QY 958 GTGCAGAACGCCAACCCGACTGCAAGACCATCTCTGGCGGCTCTCTGGCGCCCGCGGCGCACC 1017
Db 967 GTGCAGAACGCCAACCCGACTGCAAGACCATCTCTGGCGGCTCTCTGGCGCCCGCGGCGCACC 1026
QY 1018 CTGGAGGAGATGATGACCGCTGCCAGGGCGTGGGGCGGCGCCGCCACAAAGGCCCGGCTG 1077
Db 1027 CTGGAGGAGATGATGACCGCTGCCAGGGCGTGGGGCGGCGCCGCCACAAAGGCCCGGCTG 1086
QY 1078 CTGGCGGAGGCGATGAGCCAGGCCAACACAGCGTGAACATCATGATGACAGAGAGCAATTC 1137
Db 1087 CTGGCGGAGGCGATGAGCCAGGCCAACACAGCGTGAACATCATGATGACAGAGAGCAATTC 1143
QY 1138 AAGGGCCCGCGCGGCGCAAGCTCAAGTGTCTCAACTGCGCAAGAGGAGGCCACATCGCCAAG 1197
Db 1144 AAGGGCCCGCGCGGCGCATCTCAAGTGTCTCAACTGCGCAAGAGGAGGCCACATCGCCCGC 1203
QY 1198 AACTGCCGCGCGCGCGCAAGAGGCTGCTGGAAGTGGGCAAGAGGCGGCAAGGCGCACAGATG 1257
Db 1204 AACTGCCGCGCGCGCGCAAGAGGCTGCTGGAAGTGGGCAAGAGGCGGCGCACAGATG 1263
QY 1258 AAGGACTGCACCGAGCGCGGCGCAAGTCTCTGGCAAGATCTGGCCAGAGTCTGGCCAGCGGCG 1317
Db 1264 AAGGACTGCACCGAGCGCGGCGCAAGTCTCTGGCAAGATCTGGCCAGAGTCTGGCCAGCGGCG 1323
QY 1318 CGCCCCGGCAACTCTCTGCGAGAACCGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1377
Db 1324 CGCCCCGGCAACTCTCTGCGAGAGCGG-----CGCGGAGCGCGCGCGCGCGCGCGCGCG 1365
QY 1378 CCCCCCGCGGAGAGCTTCGGCTTCGAGGAGACCAACCCCGCGCGCGCGCGCGCGCGCGCGCG 1437
Db 1366 CCCCCCGCGGAGAGCTTCGGCTTCGAGGAGACCAACCCCGCGCGCGCGCGCGCGCGCGCGCG 1425
QY 1438 GACCGCGGAGCGCTACCGCGGAGCGCGCTGACCGCGCGCTGCGCAGCGCTGTTGCGCAGCGCG 1497
Db 1426 G-----ACCGGAGACCGCTGACCGCGCGCTGAGAGAGCGCTGTTGCGCAGCGCGCGCG 1473
QY 1498 CTGAGCCCACTAA 1509
Db 1474 CTGAGCCCAAGAA 1485
```

```

RESULT 14
US-10-190-435-10
; Sequence 10, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGERDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrella J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: P18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 3930
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GagCompPolmutatt_C
US-10-190-435-10
```

```

Query Match      84.4%; Score 1273.6; DB 12; Length 3930;
Best Local Similarity 92.4%; Pred. No. 1.le-251;
Matches 1397; Conservative 0; Mismatches 79; Indels 36; Gaps 4;

QY 1 ATGGCGCCCGCGCGGAGAGCTCTCTGGCGGCGGAGAGCTGGACAAAGTGGGAGAGATCCGC 60
Db 7 ATGGCGCCCGCGCGGAGAGCTCTCTGGCGGCGGAGAGCTGGACAAAGTGGGAGAGATCCGC 66
```

QY	61	CTGCGCCCCGGGCAAGBAGCACTACATGCTGAAGCACCCTGGTGTGGGCCAGCCCGGAG	120
DB	67	CTGCGCCCCGGGCAAGBAGTGCTACATGATGAAGCACCCTGGTGTGGGCCAGCCCGGAG	126
QY	121	CTGGAGGGCTTCGCCCTGACCCCCCGCCCTGCTGGAGACCCCGAGGGCTGCAAGCAGATC	180
DB	127	CTGGAGAGTTGCCCCCTGACCCCCCGCCCTGCTGGAGACCCAGCGAGGGCTGCAAGCAGATC	186
QY	181	ATGAAGCAGCTGCAGCCCCGCCCTGCAGACCGGGCACCCAGGAGCTGCCGAGCCCTGTACAC	240
DB	187	ATCCGCCAGTGCACCCCCGCCCTGCAGACCGGGCACCCAGGAGCTGGAAGAGCCTGTTCAAC	246
QY	241	ACCGTGGCCACCCCTGTACTGCTGCACGCCCGGCATCGAGGTCCGCGACACCAAGAGGCC	300
DB	247	ACCGTGGCCACCCCTGTACTGCTGCACGAGAAGATCGAGGTCCGCGACACCAAGAGGCC	306
QY	301	CTGGACAAGATCGAGGAGGAGCAGAACAACTCCCGAGCAGAAGACCCAGCAGGCCAAGGAG	360
DB	307	CTGGACAAGATCGAGGAGGAGCAGAACAACTCCCGAGCAGAAGATCCAGCAGGCCGAGGCC	366
QY	361	GCCGAC--GGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACTCGAGGGGCCAGATG	417
DB	367	GCCGACAAAGGTGAGCCAGAACTACCCCATCGTGCAGAACTCGAGGGGCCAGATG	426
QY	418	GTGCACCAAGCATCAGCCCCCGCACCGCTGAACGCCCTGGTGAAGTGATCGAGGAGAG	477
DB	427	GTGCACCAAGCATCAGCCCCCGCACCGCTGAACGCCCTGGTGAAGTGATCGAGGAGAG	486
QY	478	GCCTTCAGCCCCGAGGTGATCCCCCACTGTTACCGCGCCCTGAGCGAGGGCCACCCCGCAG	537
DB	487	GCCTTCAGCCCCGAGGTGATCCCCCACTGTTACCGCGCCCTGAGCGAGGGCCACCCCGCAG	546
QY	538	GACCTGAACACCATGTTGAACACCCGTGGGGGCCACCCAGCGCCCATGACAGATGCTGAAG	597
DB	547	GACCTGAACACCATGTTGAACACCCGTGGGGGCCACCCAGCGCCCATGACAGATGCTGAAG	606
QY	598	GACACCATCAACGAGGAGGCGCGCCGAGTGGGACCGCCCTGCACCCCGCTGCAGGCCGCCCG	657
DB	607	GACACCATCAACGAGGAGGCGCGCCGAGTGGGACCGCCCTGCACCCCGCTGCAGGCCGCCCG	666
QY	658	GTGGCCCCCGGACAGATGCGGACACCCCGCGGACGAGACATCGCGCGGACACCCAGAAC	717
DB	667	ATCGCCCCCGGACAGATGCGGACACCCCGCGGACGAGACATCGCGCGGACACCCAGAAC	726
QY	718	CTCAGGAGCAGATCGCCTGGATGACCAAGACCCCGCGTGCCTGCGGAGACATCTAC	777
DB	727	CTCAGGAGCAGATCGCCTGGATGACCAAGACCCCGCGTGCCTGCGGAGACATCTAC	786
QY	778	AAGCGGTGGATCATCCTGGGCTGAACAAGATCGTGGGATGTACAGCCCGTGAGCATC	837
DB	787	AAGCGGTGGATCATCCTGGGCTGAACAAGATCGTGGGATGTACAGCCCGTGAGCATC	846
QY	838	CTGGACATCCGCCAGGGCCCCAAGAGCCCTTCGCGACATACGTGGACCCGTTCTTCAAG	897
DB	847	CTGGACATCAAGCAGGGCCCCAAGAGCCCTTCGCGACATACGTGGACCCGTTCTTCAAG	906
QY	898	ACCTTCGCGCGCAGCAGGCCACCCAGGACCTGAAGAATCTGATACCCGAGACCCCTGCTG	957
DB	907	ACCTTCGCGCGCAGCAGGCCACCCAGGACCTGAAGAATCTGATACCCGAGACCCCTGCTG	966
QY	958	GTGCAGAACGCCAACCCCGACTGCAAGACCATCTCGCGCGCTCTCGGCCCGCGGCCAC	1017
DB	967	GTGCAGAACGCCAACCCCGACTGCAAGACCATCTCGCGCGCTCTCGGCCCGCGGCCAC	1026
QY	1018	CTGGAGGAGATGATACCCGCTGCCAGGGCGCGCGCGGCCGCCACAGGCCCGCGCTG	1077
DB	1027	CTGGAGGAGATGATACCCGCTGCCAGGGCGCGCGCGGCCGCCACAGGCCCGCGCTG	1086
QY	1078	CTGGCCGAGGCGATGAGCCAGGCCAACAGCGGTGAACATCATGATCGAGAGAGCAACTTC	1137
DB	1087	CTGGCCGAGGCGATGAGCCAGGCCAACACACCC--AGCGTGTGATCGAGAGAGCAACTTC	1146

RESULTS

US-10-19C-435-11

03 10 150 433 11 : Sequence 11. Application US/10190435

Sequence ID, Application 03/10/11
Publication No. US20030143248A1

GENERAL INFORMATION:

APPLICANT: ZUK MEGEDE, Jan

APPLICANT: BARNETT, SUSAN W.

APPLICANT: LIAN, Ying

APPLICANT: DIAN, LINDY
APPLICANT: ENGEL BRECHT, SUSAN

APPLICANT: ENGELBRECHT, SUSAN
APPLICANT: VAN BENSBUURG, ESTRELLA J.

APPLICANT: VAN KENSBURG, ESTHERICA J.

FILE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE
FILE OF INVENTION: POLYPEPTIDES POLYPEPTIDES AND USES THEREOF
FILE OF INVENTION: POLYPEPTIDES POLYPEPTIDES AND USES THEREOF

; IITE OF INVENTION: POLYPEPTIDES, POLYP
; FIVE REFERENCE: BP18133 003 / 2303-18133

FILE REFERENCE: PP18133.003 / 2302-18133
CURRENT APPLICATION NUMBER: US/10/190 435

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

CURRENT FILING DATE: 200
NUMBER OF SEC 7D NOS: 219

NUMBER OF SEC ID NOS: 316

SOFTWARE:

SEQ IS NO II

; LENGTH: 3

TYPE: DNA

; ORGANISM

FEATURE:

OTHER INFORMATION

Query Match	84.4%	Score 1273.6;	DB 12;	Length 3930;
Best Local Similarity	92.4%;	Pred. No. 1.1e-251;		
Matches 1397;	Conservative	0;	Mismatches 79;	Indels 36; Gaps 4;

QY	1	ATGGGGCCCCGGCCAGCATCTCTGGCGGGCGAAGCTGGACAACTGGGAGAGATCCGC	60
Dh	7	ATGGGGCCCCGGCCAGCATCTCTGGCGGGCGAAGCTGGACGCTGGAGCGCATCCGC	65
QY	61	CTGGCCCCCGGGCGCAAGAAGCACTACATGCTCAAGCACCTGGTGTGGGCCAGCGCCGAG	120
Dh	67	CTGGCCCCCGGGCGCAAGAAGTGTCTACATGATCAAGCACCTGGTGTGGGCCAGCGCCGAG	126
QY	121	CTGGAGGGCTTCGCCCTGAACCCCGGCCCTGCTGGAGAGACCGCGAGGGCTGCAAGCAGATC	180
Dh	127	CTGGAGAAGTTCGCCCTGAACCCCGGCCCTGCTGGAGAGACCGAGAGGGCTGCAAGCAGATC	186
QY	181	ATGAAGCAGCTGCAGCGCCGCCCTGCAGACCGGACCGAGGAGCTGCGCAGCCTGTACAAC	240
Dh	187	ATCCGCCAGCTGCACCCCGCCCTGCAGACCGGACCGAGGAGCTGAAGAGCTGTGTTCAAC	246

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2003, 17:33:23 ; Search time 97.4689 Seconds
(without alignments)
5833.434 Million cell updates/sec

Title: US-09-475-704A-4
Perfect score: 1509
Sequence: 1 atgggagccgcgcgcacat.....gcggccccctgagccagtaa 1509

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PT05_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	929.4	61.6	4307	4	US-09-552-950-2	Sequence 2, Appli
2	925.4	61.3	9772	4	US-09-552-950-5	Sequence 5, Appli
3	734.2	48.7	1496	4	US-09-184-418C-82	Sequence 82, Appl
4	734.2	48.7	8972	4	US-09-184-418C-9	Sequence 9, Appli
5	709.6	47.0	1485	4	US-09-184-418C-73	Sequence 73, Appl
6	709.6	47.0	9010	4	US-09-184-418C-8	Sequence 8, Appli
7	708	46.9	1476	4	US-09-184-418C-100	Sequence 100, App
8	708	46.9	8959	4	US-09-184-418C-11	Sequence 11, Appl
9	663.6	44.0	1486	4	US-09-184-418C-38	Sequence 38, Appl
10	663.6	44.0	8992	4	US-09-184-418C-4	Sequence 4, Appli
11	651.8	43.2	8968	4	US-09-184-418C-1	Sequence 1, Appli
12	649.4	43.0	1479	4	US-09-184-418C-12	Sequence 12, Appl
13	642.2	42.6	7399	2	US-08-418-848A-9	Sequence 9, Appli
14	642.2	42.6	9709	2	US-08-188-583-5	Sequence 5, Appli
15	642.2	42.6	9709	3	US-08-388-353-1	Sequence 1, Appli
16	642.2	42.6	9709	3	US-08-488-551B-1	Sequence 1, Appli
17	642.2	42.6	9709	4	US-09-309-572-15	Sequence 15, Appl
18	642.2	42.6	12479	4	US-09-318-138-13	Sequence 13, Appl
19	642.2	42.6	12494	3	US-08-935-312-13	Sequence 13, App
20	642.2	42.6	12494	3	US-08-848-760B-33	Sequence 33, App
21	642.2	42.6	15581	3	US-08-646-538-35	Sequence 35, Appl
22	642.2	42.6	15581	3	US-09-503-222-35	Sequence 35, Appl
23	633	41.9	9737	2	US-08-944-449-7	Sequence 7, Appli
24	633	41.9	9737	4	US-09-353-362-7	Sequence 7, Appli
25	631.4	41.8	5362	3	US-08-463-210-5	Sequence 5, Appli
26	631.4	41.8	8932	3	US-09-124-900-1	Sequence 1, Appli
27	631.4	41.8	8933	3	US-08-463-210-4	Sequence 4, Appli

23	631.4	41.8	8933	3	US-09-620-958A-3	Sequence 3, Appli
29	631.4	41.8	8933	3	US-09-620-958A-4	Sequence 4, Appli
30	631.4	41.8	8933	3	US-09-620-958A-9	Sequence 9, Appli
31	627.6	41.6	4307	4	US-09-552-950-1	Sequence 1, Appli
32	627.6	41.6	9719	4	US-09-700-304-1	Sequence 1, Appli
33	626.2	41.5	3807	2	US-08-417-210A-78	Sequence 78, Appl
34	625.8	41.5	1503	4	US-09-184-418C-64	Sequence 64, Appl
35	625.8	41.5	9050	4	US-09-184-418C-7	Sequence 7, Appli
36	625.6	41.5	9207	3	US-08-388-353-800	Sequence 800, App
37	625.6	41.5	9207	3	US-08-488-552B-800	Sequence 800, App
38	625	41.4	1503	4	US-09-184-418C-29	Sequence 29, Appl
39	625	41.4	8953	4	US-09-184-418C-3	Sequence 3, Appli
40	619.4	41.0	8987	4	US-09-184-418C-2	Sequence 2, Appli
41	615.8	40.8	1494	4	US-09-184-418C-47	Sequence 47, Appl
42	615.8	40.8	4307	4	US-09-184-418C-48	Sequence 48, Appl
43	615.8	40.8	8966	4	US-09-184-418C-5	Sequence 5, Appli
44	614.4	40.7	2095	1	US-06-333-901-5	Sequence 5, Appli
45	614.4	40.7	2095	1	US-08-456-582-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-09-552-950-2
; Sequence 2, Application: US/09552950
; Patent No. 6541248
; GENERAL INFORMATION:
; APPLICANT: Oxford Biomedica (UK) Limited
; TITLE OF INVENTION: Anti-Viral Vectors
; FILE REFERENCE: 674524-2004
; CURRENT APPLICATION NUMBER: US/09/552,950
; CURRENT FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4307
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:gagpoi-SYNap - codon
; OTHER INFORMATION: optimised gagpol sequence
US-09-552-950-2

Query Match	61.6%	Score 929.4	DB 4	Length 4307
Best Local Similarity	77.7%	Pred. No. 1.8e-128		
Matches 1173; Conservative	0	Mismatches 306	Indels 30	Gaps 3
Q7	1	ATGGGGCCCCGGCCAGCATCTCTGCGGGCGGAGAGAGCTGGACAAAGTGGGAGAAAGATCCGC	60	
Pb	1	ATGGGGCCCCGGCCAGCGTCTCTGCGGGCGGAGAGCTGGACCGCTGGGAGAAAGATCCGC	60	
Q7	61	CTGGGGCCCCGGCGGCAAGAGCACTACATGCTGAAGCAACCTGGTGTGGGCCAGCCCGGAG	120	
Pb	61	CTGGGGCCCCGGCGGCAAAAGAAAGTACAAAGCTGAAGCAACATCTGTGGGCCAGCCCGGAA	120	
Q7	121	CTGGAGGCGCTTCGCCCTGAACCCCGGGCTGCTGGAGACAAAGCGAGGGCTGCAAGCAGATC	180	
Pb	121	CTGGAGGCGCTTCGCCCTGAACCCCGGGCTCTCTGGAGACCGAGGAGGGTGTGGCGCCAGATC	180	
Q7	181	ATGAAGCAGCTGCAGCCCGCCCTGCAGACCGCGCACGAGGAGCTGCCAGCCCTGTACAAC	240	
Pb	181	CTCGGCCAACTGCAGCCCGCCCTGCAAACCGGACGCGAGGAGCTGCCAGCCCTGTACAAC	240	
Q7	241	ACCGTGGCCACCCTGTACTGGTGGCACCGCCCGGCATCGAGGTCCGCGACACCAAGGAGGCC	300	
Pb	241	ACCGTGGCCACGCTGTACTGGCTCCACCCAGCGCATCGAAATCAAGGATACGAAAGAGGCC	300	
Q7	301	CTGGACAAGATCGAGGAGGAGCAGACAAGTCCCGACGAGAGACCCAGCAGGCCCAAGGAG	360	
Pb	301	CTGGATAAATCGAAGAGGGAACAGAAATAGAGCAAAAGAGGCCCAACAGGCCCGCCGCG	360	
Q7	361	GCCGACGG-----CAAGGTGAGCCAGAACTACCCCATCGTGTGAGAACTCTGCAGAGGC	411	

```

Db      361  GACACCGGACACAGCAACCAAGTCAAGCCAGCAACTACCCCATCGGTCAGAAACATCCAGGGG 420
QY      412  CAGATGGTGCACACAGCCATCAGCCCGCCGACCCCTGAACCGCTGGTGAAGTGTATCGAG 471
Db      421  CAGATGGTGCACACAGCCATCCTCCCGCCGACCGCTGAACCGCTGGTGAAGTGGTGGAA 480
QY      472  GAGAAAGGCTTCAGCCCGCCGAGGTATCCCATGTTTACCCGCGCTGAGCGAGGGCGCCACC 531
Db      481  GAGAAAGGCTTTAGCCCGGAGGTATACCCCATGTTTCTCAGCCCTGTGAGAGGGAGCCACC 540
QY      532  CCCCAGGACCTGAACACGATGTTGAACACCGTGGGGGCCACACAGCCGCGCATGCAAGT 591
Db      541  CCCCAGATCTGAACACCATGCTCAACACAGTGGGGGACACAGCCGCGCATGCAAGT 600
QY      592  CTGAAGGACACATCAACAGAGAGCCCGCGAGTGGGACCGCTGCACCCCGTGCAGGCC 651
Db      601  CTGAAGGAGACCATCAATGAGAGGCTGCCGAATGGGATCGTGTGCATCCGTTGCACGCA 660
QY      652  GGGCCCGTGGCCCGCCGCGAGATGCGGACCGCCCGCGGACGACATCGCCGGCGCCACC 711
Db      661  GGGCCCATCGCACCGCGCGAGATGCTGAGCACGGGCTCAGACATCGCCGGAACGACT 720
QY      712  AGCACCCCTGCAGGACAGATCGCCCTGGATGACAGCAACCGCCCGCTGCGCGTGGCGGAC 771
Db      721  AGTACCCCTTCAAGAACAGATCGGCTGGATGATGACCAACNACCCACCATCCCGTGGAGAA 780
QY      772  ATCTACAAGCGGTGGATCATCTCTGGGCTGAACAAAGATCGTGGGATGTACAGCCCGCTG 831
Db      781  ATCTACAACCGTGGATCATCTCTGGGCTGAACAAAGATCGTGGGATGTATAGCCCTAAC 840
QY      832  AGCATCCTGGACATCCCGCAGGGCCCCAAGAGCCCTTCGCGACTACGTGACCCGCTT 891
Db      841  AGCATCCTGGACATCCCGCAGGGCCCCAAGAGCCCTTTCGCGACTACGTGACCCGCTT 900
QY      892  TTCAAAGACCCCTGGCGCGGAGAGAGCCACCCAGGACGCTGAAGAACTGGATGACCGAGAC 951
Db      901  TACAAAGCGCTCGCGCGGAGAGGCTAGCCAGGAGGTGAAGAACTGGATGACCGAAAC 960
QY      952  CTGCTGTGTGAGAACGCCAACCGCGACTGCAAGACCACTCTGCGCGCTCTGGCGCCCGCC 1011
Db      961  CTGCTGTGTGAGAACGCCAACCGCGACTGCAAGACCACTCTGCAAGGCGCTGGCGCGAGG 1020
QY      1012  GGCACCCCTGAGGAGATGATGACCGCGCTGCCAGGGCGTGGGGGCGCCCGGACAAAGG 1071
Db      1021  GCTACCCCTAGAGGAANTGATGACCGCGCTGTCAAGGAGTGGGAGACCGCGGACAAAGGA 1080
QY      1072  CGCGTGTGTGGCGGAGCGGATGAGCGCAGG--CGAACAGCGTGAACATCATGATGAGAA 1128
Db      1081  CGCGTGTGTGGCGGAGCGGATGAGCGCAGGTGACCAACTCCGCTACCATCATGATGAGCG 1140
QY      1129  AGCAACTTCAAGGGCGCCCGCGGCGCAACGTCAGTCTTCACTGCGGCGCAAGAGGGCCAC 1188
Db      1141  GGCAACTTTCGGAACCAACGCAAGATCGTCAAGTCTTCACTGTGGCAAGAAAGGGCAC 1200
QY      1189  ATGCGCCAGAACTGCGCGCGCCCGCGGCGCAAGAGGGCTGCTGGAAGTGGCGAAGAGGG 1248
Db      1201  ACAGCCCGCAACTGCAGGGCGCTAGGAAAGAGGGCTGCTGGAATGCGGCAAGGAAGGG 1260
QY      1249  CACAGATGAAGGACTGCACCGAGCGCCAGGCGCAACTTCTGGGCAAGATCTGGCGCCAG 1308
Db      1261  CACAGATGAAGGACTGTACGGAGAGACAGGCTAATTTTITAGGGAAGATCTGGCGCTTC 1320
QY      1309  CACAAGGGCGCGCGCGGCAACTTCTGTGAGAACCGGAGCGAGCGCGCGCGCCCGCCAC 1368
Db      1321  TACAAGGGAAGGCGAGGAATTTCTTCAGAGCAG-----ACCGCG 1362
QY      1369  CCGACCGCGCCCGCGCGGAGAGTTCGCTTCGAGGAGACCAACCCCGCGCGCGCAAGCAG 1428
Db      1363  CCAACAGCCCCACCGCGCGGAGTTCAGGTCTGGGGTGGCGGACAACTCCCGCTCG 1422
QY      1429  GAGCCCAAGGACCGGAGCCCTACCGCGAGCCCTGACCGCGCTGCGCAGCGCTGTTGCG 1488
```

```

Db      1423  AAGCAGAGCCGACCGCAGGGCAGCGGTGCTCTCAACTTCCCTCAGGTCAAGCTTTGGC 1482
QY      1489  AGCGGCCCC 1497
Db      1483  AGCGACCCC 1491

RESULT 2
US-09-552-950-5
; Sequence 5, Application US/09552950
; Patent No. 6541248
; GENERAL INFORMATION:
; APPLICANT: Oxford Biomedica (UK) Limited
; TITLE OF INVENTION: Anti-Viral Vectors
; FILE REFERENCE: 674524-2004
; CURRENT APPLICATION NUMBER: US/09/552,950
; CURRENT FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 5
; LENGTH: 9772
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: pSYNGP
US-09-552-950-5

Query Match      61.3%; Score 925.4; DB 4; Length 9772;
Best Local Similarity 77.3%; Pred. No. 6.8e-128;
Matches 1175; Conservative 0; Mismatches 316; Indels 30; Gaps 3;

QY      1 ATGGCGCGCGCGCGCGCGCATCTCGCGCGCGAGAGAGTGGCAACATCGGAGAGATCGCG 60
Db      1108 ATGGCGCGCGCGCGCGCGCATCTCGCGCGCGAGAGTGGCAACATCGGAGAGATCGCG 1167
QY      61 CTGGCGCGCGCGCGCGCGAGAGAGTCTACATGCTGAAGAGAGTGGTGGCGCGAGCGGAG 120
Db      1168 CTGGCGCGCGCGCGCGCGAGAGAGTACAAAGTGAAGAGTGAAGAGTGAAGAGTGAAG 1227
QY      121 CTGAGAGGCTTCGCGCTGAACCGCGCGCGCTGCTGGAGAGCGCGGAGGCTGCAAGCAGATC 180
Db      1228 CTGAGAGGCTTCGCGCTGAACCGCGCGCGCTGCTGGAGAGCGCGGAGGCTGCGCGCGAGATC 1287
QY      181 ATGAAGCAGCTGCAGCGCGCGCTGCAGACCGCGCGCATCGAGAGAGTGGCGAGAGCGCTGTACAAAC 240
Db      1288 CTGCGGCGCAACTGCAGCGCGCGCTGCAACCGCGCGCATCGAGAGAGTGGCGAGAGCGCTGTACAAAC 1347
QY      241 ACCGTGGCGCACCCCTGTACTGGTGCACCGCGCGCATCGAGAGTGGCGAGAGAGAGGCG 300
Db      1348 ACCGTGGCGCACCGTGTACTGGTGCACCGCGCGCATCGAGAGAGTGGCGAGAGAGAGGCG 1407
QY      301 CTGGACAGATCGAGGAGGAGCAGAAAGTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Db      1408 CTGGATAAANTCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1467
QY      361 GCGGACGG-----CAAGGTGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 411
Db      1468 GACACCGGACACAGCAACCGAGGTGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1527
QY      412 CAGATGTGCACCGAGGCGATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 471
Db      1528 CAGATGTGCACCGAGGCGATCCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1587
QY      472 GAGAAGGCTTCAGCGCGCGAGGTGATCCCATGTTTACCGCGCGCGCGCGCGCGCGCGCG 531
Db      1588 GAGAAGGCTTTAGCGCGCGAGGTGATACCATGTTTCTCAGCGCGCGCGCGCGCGCGCGCG 1647
QY      532 CCCCAGGAGCTGAACACGATGTTGAACACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCG 591
Db      1648 CCCCAGATCTGAACACCATGCTCAACACAGTGGGGGAGAGAGAGAGAGAGAGAGAGAGAG 1707
QY      592 CTGAAGGAGACCATCAACAGGAGGCGCGCGAGTGGAGAGCGCGCTGCACCGCGCTGCGAGGCC 651
```

```
Db 1706 CTGAAGGAGACCATCAATGAGGAGGCTGCCGAATGGGATCCTGTGGCATCCGCTCCACGCA 1767
QY 652 GGGCCCGGTGGCCCGCGGCGAGATCGCGGACCCCGCGGCGGCGGAGACATCGCCGGCGCCACC 711
Db 1768 GGGCCCATCGCACCGGCGCGAGATCGGTGAGCGACGGGCTCAGACATCGCCGGGAAGCACT 1827
QY 712 AGCACCCCTGCAGGAGCAGATCGCTGGATGACCAAGCAACCCCGCTGCCCGTGGGGGAC 771
Db 1828 AGTACCCCTTCAGGAACAGATCGGCTGGATGACCAACCAACCCACCCATCCCGGTGGGAGAA 1887
QY 772 ATCTACAAGCGGTGGATCACTGCTGGCCCTGACAAAGATCGTGGGGATGTACAGCCCGGCTG 831
Db 1888 ATCTACAACGCTGGATCATCTGTGGCCCTGACAAAGATCGTGGGCATGTATAGCCCTAGC 1947
QY 832 ACCATCCTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCGCGGACTACGTGGACCGCTTC 891
Db 1948 AGCATCCTGGACATCCGCCAAGGCCCGGAAGGAACCCCTTTTCGCGACTACGTGGACCGGCTTC 2007
QY 892 TTCAAGACCCCTGCGCGCGGAGAGGCCACCCAGGACGCTGACAAACTGGATGACCGGAGAC 951
Db 2008 TACAAAACGCTCCGCCGCGGAGGAGGCTAGCCAGGAGGTGAAGAACTGGATGACCGAAACC 2067
QY 952 CTGCTGGTGACAAACGCCAAGCTCCGACTGCAAGACCATCTCTGGCGCTCTGCGGCCCGCGG 1012
Db 2068 CTGCTGGTCCAGAACGCCAAGCCGGACTGCAAGACCATCTCTGAAGGCCCTTGGGCCAAG 2127
QY 1012 GCCACCCCTGAGGAGATGATCACCGCTTGCAGGGCGTGGCGGCCCGCCCGCCACAAAGGCC 1072
Db 2128 GCTACCCCTAGAGGAATGATGACCGCTTGCAGGGAGTGGCGGACCCCGCCACAAAGCCA 2187
QY 1072 CGCGTGTCTGGCGGAGGCGATGAGCCAGG---CCAACAGGTGAACATCATGATGACAGAAG 1128
Db 2188 CGCGTCTCTGGTGAAGCCATGAGCCAGGTGACCAACTCCGCTACCATCATGATGCAAGGC 2247
QY 1129 AGCAACTTCAGAGGCCCCCGCGGCAACGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCAC 1188
Db 2248 GGCAACTTTCGGAACCAACCGCAAGATCGTCAAGTGCTTCAACTGTGGCAAGAAGGGCCAC 2307
QY 1189 ATGCGCAAGAACTGCGCGCGCGCCCGCGAAGAGGGCTGCTGGAAGTGGCGAAGGAGGCC 1248
Db 2308 ACAGCCCGCAACTGCAGGGCGCCCTTAGGAAAAGGGCTGTGGAATGTGGAAAAGGAAGGA 2367
QY 1249 CACCAGATGAAGGACTGCACCGAGCGCCCGCAAGAGGGCTGCTGGAAGTGGCGAAGGAGGCC 1308
Db 2368 CACCAATGAAGATGTACTGAGAGACAGGCTAATTTTITAGGGAAGATCTGGCCCTTTC 2427
QY 1309 CACAAAGGCGCGCGCGCAACTTCTGCAAGACCGGAGCGGAGCGCGCGCGCGCGCGCGCGCG 1368
Db 2428 CACAAGGGAAGCCAGGGAATTTCTTCAGAGCAG-----ACCAGAG 2469
QY 1369 CCCACCGCGCGCGCGCGGAGAGCTTCCGCTTCGAGGAGACACCCCGCGCGCGCGCGCGCGCGCG 1428
Db 2470 CCAACAGCGCGCGCGCGGAGAGAGCTTCAGSTTTGGGGAAGAGACACAACTTCCCTCTCAG 2529
QY 1429 GAGCCCCAAGSACCGGAGCCCTACCGCGGAGCCCTGACCGCCCTTCCGCGAGCTGTTCGGC 1488
Db 2530 AAGCAGGAGCGCGATAGACAAGAACTGTATCCCTTAGCTCCCTCAGATCAGCTTTTTCGC 2589
QY 1489 AGCGGCGCGCTGAGCCAGTAA 1509
Db 2590 AGCGACCCCTCGTCACAAATAA 2610
```

RESULT 3

```
US-09-184-418C-82
; Sequence 82, Application US/09184418C
; Patent No. 6492110
; GENERAL INFORMATION:
; APPLICANT: Hahn, Beatrice
; APPLICANT: Gao, Feng
; APPLICANT: Shaw, George
; TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
; IMMUNODEFICIENCY VIRUS TYPE 1
```

```
; FILE REFERENCE: D5287
; CURRENT APPLICATION NUMBER: US/09/184.418C
; CURRENT FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 112
; SEQ ID NO 82
; LENGTH: 1496
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; OTHER INFORMATION: isolate=96ZM751.3; gene=gag
US-09-184-418C-82

Query Match 48.7%; Score 734.2; DB 4; Length 1496;
Best Local Similarity 69.0%; Pred. No. 6.9e-100;
Matches 1020; Conservative 0; Mismatches 459; Indels 1; Gaps 1;

QY 1 ATGGGCGCGCGCGCCAGCATCCCTGCGCGGCGGAGAAAGCTGGACAAAGTGGGAGAAAGATCCGC 60
Db 1 ATGGGTGCGAGAGCGGTCAATATTAAAGAGCGCGGAANTTAGATGAATGGGAAAGAATTAGG 60
QY 61 CTGGCGCGCGCGGCGCAAGAAGCACTACATGCTGAAGCAACCTGGTGTGGCGCGAGCGCGAG 120
Db 61 TTAAGGCGCGGCGGGAAGCACTATATGATGAACACTTAATATGGCGCAAGCAGGAG 120
QY 121 CTGGAGGCGCTTCGCCCTGAACCCCGGCTGCTGGAGACCGCGGAGGCTGCAAGCAATC 180
Db 121 CTGGAAGATTTGCACTTAACCCCTGGCTTTTAGAGACATCAGAAGGCTGTAACAATA 180
QY 181 ATGAAGCAGCTGCAGCGCGCCCTGCAGACCGGCACCGCAGGAGCTGCGAGCGCTGTACAAC 240
Db 181 ATACAACAGCTACACCAAGCTCTCCAGACAGGGAACAGAGGAACCTTAGTCAATATAAT 240
QY 241 ACCGTGGCCACCCCTGTACTGCTGCAGCGCGGCTCGAGGTCGCGACACCAAGGAGGCC 300
Db 241 ACAGTAGCAACTCTCTATGTGTACATGAAGAGATAAGGTACGAGACACCAAGGAAGCC 300
QY 301 CTGGACAAGATCGAGGAGGAGCAAGCAAGTCCCAGCAGAAGACCCAGCAGGCCAAGAG 360
Db 301 CTAGACAAGATAGAGGAAGCAACAAACAAAAGTCAACAAAAAANTACAAAAACAGAAAG 360
QY 361 GCCGACGCGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAAACCTGCAGGGCCAGATG 420
Db 361 ACTGGCGGAAGGTCACTCAAAATTTATCTATAGTGCAGAAATCTCCAAGGGCAATGGTA 420
QY 421 CACCAGGCCATCAGCCCGCGCACCCCTGAACGCCCTGGTGAAGGTGATCGAGGAGAAAGCC 480
Db 421 CACCAGGCTATATCACTAGAACTTTTGAATGCATGGTAAAAGTAATAGAGGAGAAAGGT 480
QY 481 TTCAGCGCGGAGGTGATCCCCCATGTTCAACCGCCCTGAGCGAGGGCGCCACCCCGCAGAC 540
Db 481 TTCACCCGAGAGGTAATACCCCATGTTTACAGCAATATCAGAAGGAGGCCACCCCAAGAT 540
QY 541 CTGAACAGATGTTGAACACCGTGGCGGCGCACCCAGCGCCCAAGCATGCTGAAGAGAC 600
Db 541 CTAAACACCATGTTAAATACAGTGGGGGACATCAAGCAGCCATGCCAAATGTTTAAAGAT 600
QY 601 ACCATCAACGAGGAGCGCGCGAGTGGGACCGCTGCACCCCGTGCAGGGCGCGCCCGCTG 660
Db 601 ACCATCAATGAGGAAGCTGCAGATGGGATAGGTTACATCCAGTACATGAGGCGCTATT 660
QY 661 GCCCGCGCGCAGATGCGCGGACCCCGCGGCGCAGCGAGATCGCGGCGCACCAAGCACCGCTG 720
Db 661 GCACCGGCCCAATAAAGAGAACCAAGGGGAAGTGACATAGCAGGAACACTACTGGTACCCCT 720
QY 721 CAGGAGCAGATCGCCTGGATGACCAAGCAACCCCGCGCTGCCCGTGGCGGACATCTACAAG 780
Db 721 CAGGAACAATAGCATGGATGACAAATAACCCACCTATTCCAGTGGGAGACATCTATAAA 780
QY 781 CGGTGATCATCCTGGGCTGAACAAAGATCGTGCAGATGACAGCCCCGTGAGGCATCCTG 840
Db 781 AGATGGATAATCTCGGGTTAAATAAAATAGTAAGAAATGTACAGCCCTGTGCAGCATCTG 840
QY 841 GACATCGCCGAGGGCCCCAAGGAGGCCCTTCCGCGACTACGTGGAGCGCTTCTTCAAGACC 900
```



```
Db 1216 GCTGAAGCAATGAGCCCAAGTAACAACATACAAATATCATGCAGAAACGACCCAGGCCA----- 1275
QY 1141 GGGGCCCCGGGCAACGTCACAGTCTTCAACTGCGGCAAGGAGGCGGCACATGCGCAAGAAG 1200
Db 1276 GGGCCCTAAAGAAATGTGTAATGTTTCACCTGCGCAGGAGGAGGCATATAGCCAGGAAT 1335
QY 1201 TGGCGGCCCCCGCCGCAAGAGGGGCTGCTGGAAATGCGGCAAGGAGGCGCCACAGATGAAG 1260
Db 1336 TGCAGGGCTCC1GGGAAAGAAAGGCTGTTGGAATG1GGAAAGGAGGACACCAAAATGAAA 1395
QY 1261 GACTGCAACGAGCGCCAGGCGCAACTTCCCTGGGCAAGATCTGAGCCAGGCGCAAGGCGCGG 1320
Db 1396 GACTGTACTGAGAGACAGGCTAATTTT1AGGAAATTTTGGCTTCCAGAGAGGGGAGG 1455
QY 1321 CCGGGCAACTTCTGTCAGAACGCGAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
Db 1456 CCGGGCAACTTCTGTCAGAACGCGAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1515
QY 1381 CCGCGCGAGAGCTCCGCTTCAGAGGAGACGACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1440
Db 1516 CCAGCAGAGAGCTTCAGGTTCCAGGAGACACACCGCTGCGCGCGCGCGCGCGCGCGCGCG 1575
QY 1441 CCGGAGCGCTACCGCGAGCGCGCTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1479
Db 1576 AAGAACCCCTTAACCTGCGCTCAATCACTCTTTGGCGAGC 1614

RESULT 5
US-09-184-418C-73
; Sequence 73: Application US/09184418C
; Patent No. 6492110
; GENERAL INFORMATION:
; APPLICANT: Hahn, Beatrice
; APPLICANT: Gao, Feng
; APPLICANT: Shaw, George
; TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
; TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS TYPE 1
; FILE REFERENCE: D6287
; CURRENT APPLICATION NUMBER: US/09/184,418C
; CURRENT FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 112
; SEQ ID NO 73
; LENGTH: 1485
; TYPE: DNA
; ORGANISM: Human: immunodeficiency virus type 1
; FEATURE:
; OTHER INFORMATION: isolate-962M51.8; gene=gag
US-09-184-418C-73

Query Match 47.0%; Score 709.6; DB 4; Length 1485;
Best Local Similarity 59.2%; Pred. No. 2.7e-96;
Matches 1008; Conservative 0; Mismatches 424; Indels 24; Gaps 2;

QY 1 ATGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 10
Db 1 ATGGGCGCGAGAGCGCTCAATATTAAGAGGGGGGAAAATTAGATAAAATCGGAAAATAATTAG 10
QY 61 CTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Db 61 CTAGGGCCAGGGGAAAGAAACGCTATATGATAAACACCTAGTATGGGCAAGCAGCGGAG 120
QY 121 CTGGAGGGCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
Db 121 CTGGAAGATTTTGGCGCTTAACCGTGGCGCTTTAGAAACATCAGAAAGCGCTGTAACAAATA 180
QY 181 ATGAAGCAGCTGCAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
Db 181 ATGAACACGCTACAAACCGCGCTTTCAGACAGCAACGCGGAGGAACTTAGATCATTTATACAAC 240
QY 241 ACCGTGGCCACCGTGTACTGCGTGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
Db 241 ACAGTAGCAACTCTCTATTGTGTACATGAAGGGGTAGAGGTACGAGACACCAAGGAAGCC 300
```

```
QY 301 CTGGACAAGATCGAGGAGCAGAGCAAGTCCCGACGAGAAAGACCCAGAGGCCA----- 355
Db 301 TTAGACAGGATAGAGGAAGAACAAACAAATAATTCAGCAAAAAATACAGCAAAAAAACACAG 360
QY 356 -AGGAGGCGCAGCGGCAAGG1GAGCCAGAACTACCCCATCGTGCAGAACTGCAGGGCGCAG 414
Db 361 CAAGCGGCTGACGGAAAGGTCAGTCAAAATTA1CCTTA1AGTGCAGAAATCTCCAAGGGCAA 420
QY 415 ATGGTGCACAGGCGCATCAGCCCGCGCACCTTCAACCGCTGGTGAAGTGTATCGAGGAG 474
Db 421 ATGGTACACAGAAACTATCACCTAGAACTT1GAATGTCATGGTAAAGTAA1AGAAGAA 480
QY 475 AAGGCTTCAAGCCCGGAGGTGATCCCGCATGTTCAACCGCTGAGCGAGGCGCGCACCCCG 534
Db 481 AAAGCTTTTAGCCCGAGAGTAA1ACCCATGTTTACAGCATTA1CAGAAAGGAGCCACCCCA 540
QY 535 CAGGACCTGAACACAGATGTTGAACACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 594
Db 541 CAAGATTTAAACACCATGTTAA1ACAGTGGGGGACATCAAGCAGCGCATGCAAA1GTTA 600
QY 595 AAGGACACCATCAACGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 654
Db 601 AAAGTACTATCAATGAGGAGGCTGCAGATGGATGGATAGATTACATCCAGTGCATGCGAGG 660
QY 655 CCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 714
Db 661 CCTATTGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
QY 715 ACCCTGCAGGAGCAGATCGCCTGGATGACGAGCAACCGCGCGCGCGCGCGCGCGCGCGCG 774
Db 721 ACCCTCCAGAACAGATAGCATGGA1GACAAAGTAATCCCGCTAT1CCAGTGGGAGACATC 780
QY 775 TACAAGCGGTGATCATCCTGGCGCTGAACAAAGATCGTGGGATGTACAGCCCGCGTGAGC 834
Db 781 TATAAAGATGGATTAATTC1GGGTTAA1TAA1AATAGTAAGATGTATAGCCCTGTGAGC 840
QY 835 ATCTGTGACATCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 894
Db 841 ATTTTGGACATAAAACAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
QY 895 AAGACCCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 954
Db 901 AAACTTTAAGAGCG1GACAGGCTACACAGAGAGTAA1AAATTTGGATGACAGACACCTTG 960
QY 955 CTGGTGCAGAACGCCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1014
Db 961 TTGTTCCAAATGCAACCCAGATTGCAAGACCAATTTTAAAGCATTTAGAGCCAGGGGCT 1020
QY 1015 ACCCTGGAGGAGATGATGACCGCGCTGCCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1074
Db 1021 ACATTAGAGAAATGATGACAGCATGTCAAGGAGTGGGAGGACCTAGCCACAAAGCAAGA 1080
QY 1075 GTGCTGGCGGAGGCGATGAGCCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1134
Db 1081 GTGTTGGCTGAGGCAATGAGCCAAACAAATAGTGAACATACTGATGACAGAAAGCAAT 1140
QY 1135 TTCAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1194
Db 1141 TTTAAAGGAAA1AAAGAA1GTTAAATGTTTAACTGTTGTAAGGTAAGGAGGCGCACATGCC 1200
QY 1155 AAGAACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1254
Db 1201 AGAAATTCAGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1260
QY 1255 ATGAAGACTGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1314
Db 1261 ATGAAGACTGTACTGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1320
QY 1315 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1374
Db 1321 GGAAGGCGCGGAAATTTCTTTCAGAACAG-----GCCAGAGCCAAACA 1362
```



```
Best Local Similarity 69.2%; Pred. No. 4.4e-95;
Matches 1004; Conservative 0; Mismatches 425; Indels 21; Gaps 2;

QY 1 ATGGGCGCGCGCCAGCATCTCGCGGCGGAGAGCTGGACAAGTGGGAGNAGATCCGCG 60
DB 1111111111111111111111111111111111111111111111111111111
138 ATGGGTGGCAGAGCGTCAATATTAAAGAGGGGGAATTAGATAGATGGGAAAAAATCGG 197
QY 61 CTGCGCCCGCGCGGCAAGCACTACATGCTGAAGCACTGGTGTGGCGCAGCGCGAG 120
DB 1111111111111111111111111111111111111111111111111111111
198 TTAAGGCCAGGGGGAAGAAACATATATGATAAAACACTTAGTATGGGCAAGCAGGAG 257
QY 121 CTGAGGGCTTCGCCCTGAACCCCGGCTGCTGGAGACGCGCGAGGCTGCAAGCAGATC 180
DB 1111111111111111111111111111111111111111111111111111111
258 CTGGAAGATTTCGCTTAAACCCIGGCTTTAGAGAGCTCAGACGGATGTAAACAATA 317
QY 181 ATGAAGCACTGCAGCCCGCCCTGCAGACCGGCGCAGGAGCTGCGCAGCCCTGTACAAC 240
DB 1111111111111111111111111111111111111111111111111111111
318 ATAAACAGCTACATCCAGCTCTTAAGACAGGAACAGAGGAACITTAGTTCATTATCAAC 377
QY 241 ACCGTGGCCACCCCTGTACTGCTGCACGCGGCGCATCGAGTCCGGACACCAAGAGGCC 300
DB 1111111111111111111111111111111111111111111111111111111
378 ACAGTAGCAACCTCTATGTGTACATGCGAGGATAGAGGTACGAGACACCAAGAGGCC 437
QY 301 CTGGACAAGATCGAGGAGGAGCAGAAAGTCCAGCAGAAAGACCCAGGCGGCGCAAGGAG 360
DB 1111111111111111111111111111111111111111111111111111111
438 TTAGACAAGATAGAGGAAGAACAAACAAAAGTCAAGCAAAAATACAGCAGGCAAGAG 497
QY 361 GCCAGCGCAAGGTGAGCCAGAACTACGCCATCGTGTGAGAACTTCGAGGCGCAGATGGTG 420
DB 1111111111111111111111111111111111111111111111111111111
498 GCTGACGGAAGGTCACTCAAAAATTATCTCTATAGTGAATAATCTCAAGGGCAATGGTA 557
QY 421 CACCAGGCCATCAGCCCGCGCACCCCTGAACGCCCTGGGTGAAGGTGATGAGGAGAAAGGCC 480
DB 1111111111111111111111111111111111111111111111111111111
558 CACCAGCCCTATCACCCTAGAACCTTTGAATGCGTGGTAAAGTAAATAGAGGAGAGGCT 617
QY 481 TTCAGCCCGAGGTGATCCCATGTCACCGGCCCTGAGCGAGGGCGCACCCCGCCAGCAC 540
DB 1111111111111111111111111111111111111111111111111111111
619 TTAGCCAGAGGTAATACCCATGTTACAGCATTTATCAGAGGAGGAGCCCGCTCTGAT 677
QY 541 CTGAACACAGATGTTGAACACCGTGGGCGGCCACCGAGCGCCCATGACATGCTGAAGGAC 600
DB 1111111111111111111111111111111111111111111111111111111
678 TTAACACCATGTTAAATACAGTGGGCGGACATCAAGCAGCCATGCAAAATGTTAAAGAT 737
QY 601 ACCATCAACGAGGAGGCGCGGAGTGGAGCCCGCTGACCCCGCTGCGAGCGCGGCGCGTG 660
DB 1111111111111111111111111111111111111111111111111111111
738 ACCATCAACGAGGAGGCTGGGATGGGATAGATTATCATCCAGTACATGACAGGCGCTAAT 797
QY 661 GCCCGCGCCAGATGCGCGACCCCGCGGCGAGCGACATCGCCGCGCCACCGACCGCTG 720
DB 1111111111111111111111111111111111111111111111111111111
798 CCACCAGCCAGATGAGAGAACCAAGGGAAGTGATATAGCGGAATACATAGTACCTT 857
QY 721 CAGGAGCAATGCTGCTGATGACCAAGCAACCCCGCTGCGGATGTACAGCCCGTGACATCAAG 780
DB 1111111111111111111111111111111111111111111111111111111
858 CAGGAACAATAGCATGATGATGACAGGTAAACCCACCTATCCAGTGGGAGACATCAATAA 917
QY 781 CGGTGATCATCTCGGCTGAACAAGATCGTGGGATGTACAGCCCGTGAGCATCGTG 840
DB 1111111111111111111111111111111111111111111111111111111
918 AGATGGATAATTCTGGGTTAAATAAATPASTAAGATGTATAGCCCTGTCAGCATTTG 977
QY 841 GACATCCGCGAGGCGCCCAAGGAGCCCTTCGCGACTAGTGGACCGCTTCTTCAAGAAC 900
DB 1111111111111111111111111111111111111111111111111111111
978 GACATAAGACAAGGCGCAAGGAACCCCTTAGAGACTATGACCGGTTCTTTAAAACT 1037
QY 901 CTGCGCGCGAGCAGGCCACCCAGGAGCTGAAGAACTTGGATGACCGAGACCTTGTGGTG 960
DB 1111111111111111111111111111111111111111111111111111111
1038 TTAAGAGCTGAACAAGCTACACAAGAACTAAAAGGTTGGATGACAGACACCTTGTGGTC 1097
QY 961 CAGAACGCCAACCCCGACTGCAAGACCATCTCTGCGGCTCTCGGCGCCCGCGCGCCCTG 1020
DB 1111111111111111111111111111111111111111111111111111111
1098 CAAAATGCAAAACCCAGATGTGAAGACCATTTTAAGAGCATTAGGACCGGGCTTTCATTA 1157
QY 1021 GAGGAGATGATGACCGCTGCCAGGGGCTGGGCGGCGCCCGGCGCACAGGCGCGGCTG 1080
DB 1111111111111111111111111111111111111111111111111111111
```

```
DB 1158 GAAGAAATGGTACAGCATGTCAAGAGTGGGAGGACCTAGCCACAACCAAGAGTGTG 1217
QY 1081 CCCGAGGGGATGAGCCAGCCCAACAGCGTGAACATCATGTGACAGAGGCAACTTCAAG 1140
DB 1111111111111111111111111111111111111111111111111111111
1218 GCTGAGGCATGAGCCCAATCACATAG---TAACATAATGTGACAGAGGCAACTTTAAA 1274
QY 1141 GGCCCCCGCGCAACGTCAAGTGTCTCAACTGCGGCAAGAGGGGCCACATCGCCCAAGAAC 1200
DB 1111111111111111111111111111111111111111111111111111111
1275 GGCCCTAAAGAAATTGTTAAATGCTTCAACTGTGGCAAGGAAGGACATAGCCAGAAAT 1334
QY 1201 TGCGCGCGCCCGCGCAAGAGGGCTGCTGGAAGTGGGCAAGGAGGGCCACCAGATGAAG 1260
DB 1111111111111111111111111111111111111111111111111111111
1335 TGCAGGGCCCGCTAGAAAAGAGGGCTGTGGAAATGTGGCAAGAGGACACCAAAATGAAA 1394
QY 1261 GACTGCACCGAGCGCCAGGCCAACTTCTCTGGGCAAGATCTGGCCCGACACCAAGGCGCG 1320
DB 1111111111111111111111111111111111111111111111111111111
1395 GACGTACTGAGAGGCGAGGCTAAATTTTTCAGGAAATTTTGGCCTTCCACAAAGGGAGG 1454
QY 1321 CCCGGCAACTTCTGTCAGAACCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
DB 1111111111111111111111111111111111111111111111111111111
1455 CCAGGGAATTTCCTTCAAAACAG-----GCCAGAGCCCAACAGCCCA 1496
QY 1381 CCCGCCGAGAGCTTCGCTTCGAGGAGACACCGCCCGCGCGCGCGCGCGCGCGCGCGCG 1440
DB 1111111111111111111111111111111111111111111111111111111
1497 CCAGCAGAGAGCTTCAGGTTCAAGGAGACAAACCCCGCTCCGGAAGCAGGAGTGGAAAGAC 1556
QY 1441 CGCGAGGCCCT 1450
DB 11111111
1557 AGGGAACCCCT 1566

RESULT 9
US-09-184-418C-38
; Sequence 38, Application US/09184418C
; Patent No. 6492110
; GENERAL INFORMATION:
; APPLICANT: Hahn, Beatrice
; APPLICANT: Gao, Feng
; APPLICANT: Shaw, George
; TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
; TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS TYPE 1
; FILE REFERENCE: D6287
; CURRENT APPLICATION NUMBER: US/09/184,418C
; CURRENT FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 112
; SEQ ID NO 38
; LENGTH: 1486
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; OTHER INFORMATION: isolate=92RW009; gene=gag
US-09-184-418C-38

Query Match 44.0%; Score 663.6; DH 4; Length 1486;
Best Local Similarity 68.8%; Pred. No. 1.5e-89;
Matches 942; Conservative 0; Mismatches 424; Indels 4; Gaps 2;

QY 1 ATGGGCGCGCGCCAGCATCTCGCGGCGGAGAGCTGGACAAGTGGGAGAAAGA-TCCG 59
DB 1111111111111111111111111111111111111111111111111111111
1 ATGGGTGCGAGAGCGTCAATATTAAAGAGGCGGAAATTAGATGCTTGGGAAAAAATTAAG 60
QY 60 CTGCGCGCGCGCGCGGCAAGAGCACTACATGCTGAAGCACTGCTGTGGCGCAGCGCGGA 119
DB 1111111111111111111111111111111111111111111111111111111
61 TTAAGCCAGGGGGGAAGAAACATATATGATGAACACCTTAGTATGGCAAGCAGGGA 120
QY 120 GCTGAGGGCTTCGCCCTGAACCCCGGCTGCTGAGACCCCGGAGGCTGCAAGCAGAT 179
DB 1111111111111111111111111111111111111111111111111111111
121 GCTGGAAGATTTGCACTTAACCTGACCTTTAGAGACACAGAAAGGCTGTAAACAAT 180
QY 180 CATGAAGCAGCTGCAGCCCGCCCTGCAGACCGGCGCACCGGAGGAGCTGCCGAGCCCTGTACAA 239
DB 1111111111111111111111111111111111111111111111111111111
181 AATGAGACAGTGCACCCAGCTCTTCAGACAGGAACAGATGAACCTTAGTTCATTATAAA 240
```


QY 240 CACCGTGGCCACCCCTGCTACTGGTGGACGGCCGGCATCGAGGTCGGGACACCAAGAGGC 299
Db 241 TACAGTAGCAACCCCTCTATTGTGTACATCAAAAGATAGATGTAAAGACACCAAGAGGC 300
QY 300 CCGGACAAAGATCGAGGAGGAGGACAAAGTCCCGACAGAGACCCACGAGGCCAAGGA 359
Db 301 CTTAGACAAGATAGAGGAAGAACAAAGAGTCAAGCAAAAGACAGGAGGAGAGGC 360
QY 360 GGCCGAC--GGCAAGGTGAGCCAGAACACTACCCCACTGGTGCAGAACTGCAGGGGAGAT 416
Db 361 AGCTGACAAAGGAAAGTCAAGTCAAAATTAACCTATAGTGCATAATTCACAAGGCAAAAT 420
QY 417 GGTCACCAAGCCATCAGCCCGCCGACCCCTGAACGGCTGGTGAAGGATGATCGAGAGAA 476
Db 421 GGTACACCAAGCCATATCACCTAGAACTTTGAATGCGTGGTAAAGTAATAGAGAGAA 480
QY 477 GGCCCTCAGCCCGAGGTGATCCCATGTTTCACCGGCTGAGCGAGGGGCCACCCCA 536
Db 481 GGCTTTAGCCCAAGAGTAATACCCATGTTTACAGCATATATCAGAAGGAGCCACCCACA 540
QY 537 GGACCTGAACACGATGTTGAACACCGTGGGCGGCGCACAGGCGGCATGCGATGCTGAA 596
Db 541 AGATTTAAACACCATGCTAAATACAGTGGGGGACATCAAGCAGCCATGCAAAATGCTAAA 600
QY 597 GGACACCATCAACGAGGAGGCGCCGAGTGGGACCGCTGCACCCGCTGCAGGGCGGCG 656
Db 601 AGATACAAATCAATGAGGAGGCTGCAGAGTGGATAGGTACATCCAGTGCAGGCGAGGCC 660
QY 657 CGTGGCCCGCCGCGAGATCGCGACCCCGCGGCGACGACATCGCCGCGCCACACAGCAC 716
Db 661 TGTGGCCGAGCCAGATAAGAGAAACCAAGGGGAAGTACATAGCAGAACTACTAGTAC 720
QY 717 CCGTGCAGGAGCAGATCGGCTGGAAGACCAAGCAACCCCGCGTGCCTGGGCGACATCTA 776
Db 721 CCTCAGGAACAAATAGCATGGAAGACAAATACCCACCTATTCAGTGGGAGAAATTA 780
QY 777 CAAGCGGTGGATCATCTCGGCTGAACAAAGATCGTGGGAAGTACAGCCCGCTGAGCAT 836
Db 781 TAAAGATGGATAATTCCTGGGTTAAATAAATAGTAAGAAATGATAGCCCTGTACGAT 840
QY 837 CCTGGACATCCCGCAGGCGCCCAAGGAGCCCTTCGGGCACTACGTGGACCGCTTCTTCAA 896
Db 841 ATTGGACATAAAACAAGGCGCCAAAGGAACCTTTTAGAGACTATGTAGACCGGTCTTAA 900
QY 897 GACCCCTGGCGCGGAGAGGCGCCACCCAGGACGTGAAGAACTGGATGACCGGAGACCCGT 956
Db 901 AACCTTAAGAGCCGACAAAGCTTCACAGATGTAAATAATTTGAAGACAGATACCTTGT 960
QY 957 GGTGCAGAACGCCAACCCGACTGCACAGACCATCTCTCGCGGCTCTCGGCGCCCGCGCAC 1016
Db 961 AGTCCAAATGCGAACCCAGATGTAAAGACCAATTTAAGAGCATTTAGGCGCAGGGGCTC 1020
QY 1017 CCTGGAGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCGCCCGGCGCACAGGCGGT 1076
Db 1021 ATTAGAAAGAAATGATGACAGCATGCCAGGGAGTGGGAGGACCCCGCCATAAAGCAAGCT 1080
QY 1077 GCTGGCCGAGCGGATGAGCCAGGCCAACAGCGGTGAACATCATGATGCAAGAGCAACTT 1136
Db 1081 TTTGGCTGAAGCAATGAGCCAAAGTACAAACCAACCAATTAATGATGCAAGAGGCAATTT 1140
QY 1137 CAAGGGCCCGCGCGCAAGGTCAAGTGTCTCAACTGCGCAAGGAGGCGCACATCGCCAA 1196
Db 1141 TAAGGGCCAGAGAAGAAATTAAGTGTTCAACTGTGCAAAAGAGGACACCTAGCCAG 1200
QY 1197 GAACTGCGCGCGCCCGCGCAAGAGGGCTGCTGGAAGTGGGCAAGGAGGCGCCACAGAT 1256
Db 1201 AAATTCAGGGCCCCCTAGAAAAAGGGCTGTGGAAATGTGGAAGGAGGAGGACACCAAT 1260
QY 1257 GAAGGACTGACCGGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCAGCCACAGGG 1316
Db 1261 GAAAGACTGCACTGAGAGACAGGCTAATTTTAGGGAATTTGGCCCTTCCAAACAAGG 1320
QY 1317 CCGCCCCGGCAACTTCCTGTCAGAACCGGAGGCGCGCGCGCCCGCCCGCC 1366

Db 1321 GAGGCCAGGAAATTTCCCCAGAGCAGACTGGAGCCAAACAGCCCCACCAG 1370
RESULT 10
US-09-184-418C-4
; Sequence 4, Application US/09184418C
; Patent No. 6492110
; GENERAL INFORMATION:
; APPLICANT: Hahn, Beatrice
; APPLICANT: Gao, Feng
; APPLICANT: Shaw, George
; TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
; TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS TYPE 1
; FILE REFERENCE: D6287
; CURRENT APPLICATION NUMBER: US/09/184,418C
; CURRENT FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 112
; SEQ ID NO 4
; LENGTH: 8992
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; OTHER INFORMATION: isolate=92RW009; 139.1624:gag; 1690.4428:pol(N-terminus unce
; OTHER INFORMATION: 4375.4951:vif; 4891.5181:vpr; 5162.7801:tat; 5301.7958:rev;
; OTHER INFORMATION: 5403.5648:vpu; 5566.8148:env; 8150.8773:nf
US-09-184-418C-4
Query Match: 44.0%; Score 663.6; DB 4; Length 8992;
Best Local Similarity 68.8%; Pred. No. 1.4e-89;
Matches 942; Conservative 0; Mismatches 424; Indels 4; Gaps 2;
QY 1 ATGGGCGCCCGCCCGCAGCATCTCTCGCGGGCGAGAAAGCTGGACAAAGTGGGAGAGA-TCCG 59
Db 139 ATGGGTGCGAGAGCGTCAATATTAAAGAGGGGGGAAATTTAGATGCTGGGAAATAATTAG 198
QY 60 CCGCGCGCCCGCGGCAAGAAAGCACTACATGCTGAAGCACCTGCTGGTGGCGGCGCCGGA 119
Db 199 TTAAGCCAGGGGGGAAAGAAACATATATGATGAACACCTAGTATGGGCAAGCAGGGA 258
QY 120 GCTGGAGGGCTTCGCCCTGAACCCCGGCTGCTGGAGACCGCGGAGGCTGCAAGCAGAT 179
Db 259 GCTGAAAGATTTGCACCTTAACCCCTGACCTTTTAGAGACACCCAGAGGCTGTAAACAAAT 318
QY 180 CATGAAGCAGCTGCAGCCCGCCCTGCAGACCGCGGAGGAGTGGCAGCCCTGTACAA 239
Db 319 AATGAGACAGCTGCACACAGCCTTCAGACAGGAAACAGATGAACCTTAGGTCATTATATA 376
QY 240 CACCGTGGCCACCCCTGTACTGCTGCACCGCGGATCGAGGTCCGCGACACCAAGGAGGC 299
Db 379 TACAGTAGCAACCCCTCTATTGTGTACATCAAAAGATAGATGTAAAGACACCAAGGAGGC 438
QY 390 CCGTGCACAAGATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 359
Db 439 CTTAGACAAGATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 498
QY 360 GGCCGAC--GGCAAGGTGAGCCAGAACTACCCCACTGTCAGAACTGTCAGAGACCTGCAGGCCAGAT 416
Db 499 AGCTGACAAAGGAAAGTCAAGTCAAAATTTACCCCTATAGTGCATAATTCACAAGGCAAAAT 558
QY 417 GGTGCACCAAGCCATCAGCCCGCCGACCCCTGAACGGCTGGTGAAGGATGATCGAGAGAA 476
Db 559 GGTACACCAAGCCATATCACCTAGAACTTTGAATGCGTGGTAAAGAGTAATAGAGAGAA 618
QY 477 GGCCCTCAGCCCGGAGGTGATCCCATGTTTCACCGGCTGAGCGAGGGCGCCACCCCA 536
Db 619 GGCTTTAGCCCAAGAGGTAATACCCCATGTTTACAGCATTAATCAGAGGAGGAGGAGGAGGAG 678
QY 537 GGACCTGAACACGATGTTGAACACCGTGGGCGGCGCCACAGGCGGCCCAATGAGATGCTGAA 596
Db 679 AGATTTAAACACCCATGCTAAATACAGTGGGGGAGACATCAAGCAGGAGGAGGAGGAGGAG 738
QY 597 GGACACCAATCAACGAGGAGGCGCGCGGAGTGGGAGGCGGCTGCACCCCGCTGCAGGCGGCGCC 656

```
Db      739 AGATACAATCAATGAGGAGC1GCAGAGTGGGATAGGTACATCCAGTGCAGGAGGACC 796
Qy      657 CGTGGCCCGCGCCAGATGCGGAGACCCCGCGGAGAGATCGCGGCGCCACACAGAC 716
Db      799 TGTTCGCCAGGCCAGATAGAGAGAACCCAGGGGAAGTACATAGCAGGAACTACTATAC 858
Qy      717 CCTGCAGGAGCAGATCGCTGCGATGACCAAGCAACCCCGTGCCTGCGGCGGACATCTA 776
Db      859 CTTTCAGGAACAAATAGCATGGATGACAAATACCCACCTATTCAGAGTGGGAGAAATTA 918
Qy      777 CANGCGGTGATCACTCTGCGCTGAACAAGATCGTGGGATGTACAGCCCGGTGASCAT 836
Db      919 TAAAGATGGATAATCTGGGTTAAATAAAATAGTAAGAATGTATAGCCCTGTCAUCAT 978
Qy      837 CCTGGACATCCCGCCAGGCGCCCAAGGAGGCGCTTCCGCGACTACGTGGACCGCTTCTCAA 896
Db      979 ATTGGACATAAAACAAGGCCCAAGGAACCTTTTAGAGACTATGTAGACCGGTTCTTTAA 1038
Qy      897 GACCTTCGCGCCGAGCAGCGCCACCCAGGACCTGAAGACTGGATGACCGAGACCCCTGCT 955
Db      1039 AACCTTAAGAGCCGACAACTTCACAAGATGTAAAAATTTGATGACAGATACCTTGT 1098
Qy      957 GGTGCAGAGCCCAACCCGACTGCAAGACCACTCTGCGGCTCTCGGCCCGGCGGCAC 1016
Db      1099 AGTCCAAATGCGAACCCAGATTTGAAGACCAATTTTAAGAGCATTAGGCGCAGGGCTTC 1158
Qy      1017 CCTGGAGGAGATGATGACCGGCTGCCAGGGGCTGGGCGCCCGGCAAGGCGCGGT 1076
Db      1159 ATTGAGAAATGATGACAGCATGCCAGGAGTGGGAGCGCCGCGCATAAAGCAAGGT 1218
Qy      1077 GCTGGCCGAGCGGATGAGCCAGGCGCAACAGCGTGAACATCATGATGACAGAGCAACT 1136
Db      1219 TTTGGCTGAAGCAANTGAGCAAGTACAAACAACCAACATTAATGATGACAGAGGCAATT 1278
Qy      1137 CAAGGGCCCGCGCGCAACTCAAGTCTCAACTGCGGCAAGGAGGCGCCACATCGGCAA 1196
Db      1279 TAAGGCCACAGAGAAGAAATATTAAAGTGTTCAACTGTGGCAAGAGGACACCTAGGCG 1338
Qy      1197 GAACCTGCGCGCCCGCCCGCAAGAGGCGTCTGGAAGTGGGCAAGGAGGCGCCACACAT 1256
Db      1339 AAATTGACGGGCCCTAGAAAAAAGGCGCTTGGAAAATGTGGAAGGAGGAGGACACCAAT 1398
Qy      1257 GAAGGACTGCACCGAGCGCGGAGGCGCAACTCTCTGGCAAGATCTGGGCGAGCCACAGGG 1316
Db      1399 GAAAGACTGCCTGAGAGACAGGCTAATTTTATAGGGAATTTGGCCTTCCCAACAAGGG 1458
Qy      1317 CCGCCCGCGCAACTCTCTGCAAGACCGGAGCGGAGCGCGCGCGCCCGCCACCG 1366
Db      1459 CAGGCCAGGAATTTTCCCGAGAGCAGACTGGAGCGCAACAGCGCCCGCCACCG 1508
```

```
RESULT 11
US-09-184-418C-1
; Sequence 1, Application US/09184418C
; Patent No. 6492110
; GENERAL INFORMATION:
; APPLICANT: Hahn, Beatrice
; APPLICANT: Gao, Feng
; APPLICANT: Shaw, George
; TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
; TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS TYPE 1
; FILE REFERENCE: D6287
; CURRENT APPLICATION NUMBER: US/09/184,418C
; CURRENT FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 112
; SEQ ID NO 1
; LENGTH: 8968
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; OTHER INFORMATION: isolate=93BR020; 133..1611:gag; 4369.4947:vif;
; OTHER INFORMATION: 4887.5177:vpr; 5158.7767:tat; 5297.7957:rev;
```

```
; OTHER INFORMATION: 5396.5641:vpu; 5559.8099:env; 8101.8727:nef
US-09-184-418C-1
Query Match      43.28; Score 651.8; DB 4; Length 8968;
Best Local Similarity 65.78; Pred. No. 7.3e-88;
Matches 991; Conservative 0; Mismatches 497; Indels 21; Gaps 2;

Qy      1 ATGGGGCGCCCGCCAGCATCTCTGCGGCGGAGAAAGCTGGACAAGTGGGAGAAATCCGC 60
Db      133 ATGGGTGCGAGAGCGCTCAGTATTAAGCGGGGAAAATATAGCTTAAACAATCTAGTATGGGCAAGCAGGGAG 192
Qy      61 CTGGCGCCCGCGGCAAGAGCACTACATGCTGAAGCACCTGGTGTGGCGCAGCCCGGAG 120
Db      193 TTAAGGCGCGGGGAAAGAAAATATAGACTAAAACAATCTAGTATGGGCAAGCAGGGAG 252
Qy      121 CTGGAGGGCTTCGCCCTGAACCCCGGCTGCTGGAGACCGCGGAGGCTGCAAGCAGATC 180
Db      253 CTAGAACGATTGCACTTGATCCAGGCTTCTTAGAACAATCAGAAGGCTGTCGAAAAATA 312
Qy      181 ATGAAGCAGCTGCAGCCCGCCCTGCAGACCGGCAAGGAGCTGCGCAGCTGTACAAC 240
Db      313 ATAGGACAGTTACAACCATCCCTTCAGACAGGATCAGAAGAGCTCAAAATCATTAATAAT 372
Qy      241 ACCGTGGCCACCTGTACTGCGTGCACGCGCGGCTCGAGGTCGCGGACACCAAGGAGGC 300
Db      373 ACAATAGCAGTCTCTATATGTACATCAAAAGTGAAGGTAAAAGACACCAAGGAGGCT 432
Qy      301 CTGCACAAGATCGAGGAGGAGAGAAACAAGTCCACAGCAAGACCCAGCAGGCAAGGAG 360
Db      433 TTAGAGAAGCTAGAGGAAGAAACAAAACAAAGTCGGGCAAAAGACACAGCAGGACTGCT 492
Qy      361 GCGACGCGCAAGGTGAGCCAGAACTACCCCATCTGTCAGAACTGTCAGGGCCAGATGGTG 420
Db      493 GAAAAAGG--GGTCAGTCAAAATACCCCTATAGTACAGAACTTCAGGGACAAAATGGTA 549
Qy      421 CACAGGCGCATCAGCCCGCGACCCCTGAACGCTGGTGAAGTGTATCGAGGAGAAAGGC 480
Db      550 CACAGTCTTATCACCTAGAACTTTAATGTGATGGTAAAGTGTATAGAGAGAAAGGCT 609
Qy      481 TTCAGCCCGGAGGTGATCCCATGTTTACCGCGCTGAGCGAGGCGCCACCGCCAGGAC 540
Db      610 TTTAGTCCAGAGTAATACCCATCTTTTICAGCATTTATCAGAGGGGCGCCACTCCACAAGAT 669
Qy      541 CTGAACACGATGTTGAACACCGTGGGCGGCGCCAGCGCCGCTGATGATGCTGAAGGAC 600
Db      670 TTAACACACCATGTTAAATACAGTGGGGGACATCAAGCAGCGCATGCATAATGTTAAAGAG 729
Qy      601 ACCATCAACGAGGAGCGCGCGAGTGGACCGCTGCACCCGTCAGGCGCGGCCCGCTG 660
Db      730 ACCATCAATGAGGAGGCTGCAGAAATGGACAGATTACATCCAAACAGGACGACCCCATC 789
Qy      661 GCGCCCGGCGCAGATGCGCGACCCCGCGGAGGAGACATCGCGGCGCCACCGACCCCTG 720
Db      790 CCGCCAGGTCAGATAAGGGAACCTTAGGGGAAGTGTATAGCTGGAACCTACTAGTACCCCTI 849
Qy      721 CAGAGCAGATCGCTTGGATGACCAAGCAACCCCGCGCTGCCGTGGCGGACATCTACAAG 780
Db      850 CAGGAACAAAATACAATGGATGACAGGCAACCCACTGTCCCAGTGGGAGAAATGTATAAA 909
Qy      781 CGTGGATCATCTGGGCGCTGAACAAGATCGTGGCGATGTACAGCCCGCTGAGCATCTCIG 840
Db      910 AGATGGATCATCTTAGGATTAATAAAATAGTAAGAATGTATAGCCCTGTCCGCATTTTG 969
Qy      841 GACATCCCGCAGGCGCCCAAGGAGCCCTTCCCGCACTACGTGGACCGCTTCTTCAAGACC 900
Db      970 GACATAAGACAAAGGCGCCAAAAGAACCCCTTTTAGAGACTAATGTAGACAGGTTCTTTAAAC 1029
Qy      901 CTGCGCGCGGAGCAGGCGCCACCGAGGACGCTGAAGAACTGGATGACCGGACCGCTGCTG 960
Db      1030 CTAAGAGCTGAGCAAGCTACACAGGAAGTAAAGGTTGGATGACAGACACCTTGTGTC 1089
Qy      961 CAGAACGCCAACCCGACTGCAAGACCATCTCGCGCTCTCGGCCCGCGGCCACCCCTG 1020
```

```
Db 1090 CAAATGCGAACCAGATTGTAGACACATTITAAAGCATTGGGACCAGGGGCIACACTA 1149
QY 1021 GAGGAGATGATGACCGGCTGACAGGCGGTGGGCGGCCCGGCCACAAAGCGCGCGTGTG 1080
Db 1150 GAGGAATGATGACAGCATGTGAGGGAGTGGGAGGACCTAGGCATAGGCAAGATTTT 1209
QY 1081 GCGGAGGGGATGAGCCAGGCGCAACAGCGTGACATCATATGATGAGAGAGAGAACTTCA 1140
Db 1210 GCTGAGGCAATGAGCCAAAGCAACAATACAGCTATATATGATGCAAGAAAGTAACTT 1269
QY 1141 GGCCCCGGCGGCAAGCTCAAGTCTTCAACTGCGGCAAGGAGGGGCCACATCGCCCAAGA 1200
Db 1270 GCGCAAGAGAAATGTTAAATGCTTAAATGCTGCGCAAGAGAGGACACATAGCCAAA 1329
QY 1201 TGCGGCGCCCCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGGAGGGGCCACATCGCC 1260
Db 1330 TGCAGGGCCCCCTAGAAAAGAGGCTGTTGGAAGTGTGGAAGAGGAGGACACACCAAT 1389
QY 1261 GACTGCACCGAGCGCCAGGCCAACTTCTCTGGCAAGATCTGGCCGAGCCACAAAGGCG 1320
Db 1390 GACTGCACCTGAGAGACAGGCTAATTTTAAAGGAAATTTGGGCTTCCACAGGGGAG 1449
QY 1321 CCGGGCAACTTCTGTCAGAACCGCAGCGAGCGCGCGCGCCCGCCAGCTGTCACCGCC 1380
Db 1450 CCGGGAACCTTCATCCAGAACAG-----GCGAGAGCGGTGAGCGCGG 1491
QY 1381 CCGGCGGAGAGCTTCGCTTCGAGGAGACCAACCGCGCGCGCGCGCAAGGAGGCGCA 1440
Db 1492 CCAGCAGAGAGCTTCAGGTTCCGGGAGGAGACAAACCCCATCTCCGAAGCAGGAGCA 1551
QY 1441 CCGGAGCCCTACCGGAGCCCTGACCGCGCTGCGCAGCGCTGTCGAGCGGCGCGCGCT 1500
Db 1552 GACGAGGAGCTGACCTCCCTTAGCTTCCCTCAAAATCACTCTTGGCAACGACCCCTAG 1611
QY 1501 AGCCAGTAA 1509
Db 1612 TCACAATAA 1620

RESULT 12
US-09-184-418C-12
; Sequence 12, Application US/09184418C
; Patent No. 6492110
; GENERAL INFORMATION:
; APPLICANT: Hahn, Beatrice
; APPLICANT: Gao, Feng
; APPLICANT: Shaw, George
; TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
; TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS TYPE 1
; FILE REFERENCE: D6287
; CURRENT APPLICATION NUMBER: US/09/184.418C
; CURRENT FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 112
; SEQ ID NO 12
; LENGTH: 1479
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; OTHER INFORMATION: isolate=93BR020; gene=gag
US-09-184-418C-12

Query Match 43.0%; Score 649.4; DB 4; Length 1479;
Best Local Similarity 65.8%; Pred. No. 1.8e-87;
Matches 985; Conservative 0; Mismatches 491; Indels 21; Gaps 2;

QY 1 ATGGGCGCGCGCGCAGCATCTCGCGCGCGGCGGAGAGCTGGACAAGTGGGAGAGATCGCG 60
Db 1 ATGGGTGCGAGAGCGTCAGTATTAAAGCGGGGAAATATAGATGCTTGGGAAAAAATTCGG 60

QY 61 CTGCGCGCGCGCGCGCAAGCAAGCACTACATGCTGAAGCACCTGCTGGCGCGCGGAG 120
Db 61 TTAAGCGCGCGGGAAGAAAATAATAGACTAAACATCTACTATATGCTGGCAAGCAGGAG 120
```

```
QY 121 CTGGAGGCGCTTCGCGCTGAACCCCGCGCTGCTGGAGACCGCGGAGGCTGCAAGGAGATC 180
Db 121 CTAGAACGATTTGCACCTTGATCCAGGCGCTTCTAGAAACATCAGAAAGGCTGTCGAAAA 180
QY 181 ATGAACGAGCTGCAGCGCGCGCTGTCAGACCGGCGGCAAGGAGTGGCGAGCCCTGTACA 240
Db 181 ATAGGACAGTTACAACCATCCCTTCAGACAGGATCAGAAAGGCTCAANTCATATATAAT 240
QY 241 ACCGTGGCCACCTGTACTGCGTGCAGCGCGGCTGAGGTCGCGGACACCAAGGAGGCC 300
Db 241 ACAATAGCAGTCCCTCTATATGATCATCAAAAGGTAGAGGTAAAGACACCAAGGAGGC 300
QY 301 CTGGACAAGATCGAGGAGGAGCAGAACAGTCCCAGCAGAGACCCAGCAGGCGCAAGGAG 360
Db 301 ITAGAGAAGCTAGAGGAAGAACAAACAAAGGTGGCAAAAGACACAGCAAGCGACTGCT 360
QY 361 GCGGACGCGAAGTGCAGCCAGAACTACCCCATCTGTCAGAAACCTGAGGAGGCGCGAGAT 420
Db 361 GAAAAGG---GGTCAGTCAAAATTAACCTATAGTACAGATCTTCAGGGACAAATGGTA 417
QY 421 CACGAGGCAATCAGCCCGCGCACCTGAAACGCGCTGGGTGAAGGTGATCGAGGAGAGGCC 480
Db 419 CACGAGTCTTTATACCTAGAACTTTAAATGCATGGGTAAAGTGTATAGAGAGAAAGCT 477
QY 481 TTAGCGCGGAGGTGATCCCATGTCACCGCGCTGAGGAGGCGCGGCGGCGGCGGCGGAG 540
Db 478 TTTAGTCCAGAAGTAAATACCATGTTTTCAGCATTTATCAGAAAGGCGGCGGCGGCGG 537
QY 541 CTGAACAGCATGTTGAACACCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
Db 538 TTAACACCAATGTTAAATACAGTGGGCGGACATCAAGAGCGCATGCAAAATGTTAAAG 597
QY 601 ACCATCAAGGAGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 660
Db 598 ACCATCAATGAGGAGGCGTGCAGAAATGGACAGATTACATCAACACAGGCGGCGGCGG 657
QY 661 GCGCGCGCGCGAGATGCGGCGACCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 720
Db 658 CCGCGAGGTCAGATAAGGGAACCTAGGGAAGTGTATAGTGGAACTACTAGTACCGCTT 717
QY 721 CAGGAGCAGATCGCTGATGACCAAGCAACCCCGCGGCGGCGGCGGCGGCGGCGGCGGCG 780
Db 718 CAGGAACAATAACAATGGATGACAGGCGCAACCCCGGCGGCGGCGGCGGCGGCGGCGG 777
QY 781 CGGTGGATCATCTCGGCGTGAACAAGATCGTGGGATGTACAGCGCGGCGGCGGCGGCGG 840
Db 778 AGATGGATCATCTAGGATTAATAAATAAGTAAAGTATAGGCGGCGGCGGCGGCGGCGG 837
QY 841 GACATCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
Db 838 GACATAAGACAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 897
QY 901 CTGGCGCGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 960
Db 898 CTAGAGCTGAGCAAGCTACACAGGAAGTAAAGGTTGGATGACAGACACCTTGTGGTC 957
QY 961 CAGAACGCGCAACCCCGGCGTCAAGACCATCTCGCGCGGCGGCGGCGGCGGCGGCGGCG 1020
Db 959 CAAATGCGAACCCAGATGTAAAGCCATTTTAAAGCATTTGGGCGGCGGCGGCGGCGGCG 1017
QY 1021 GAGGAGATGATGACCGCGCTGCCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1080
Db 1018 GAGGAATCATGACAGCATGTGAGGAGTGGGAGGCGGCGGCGGCGGCGGCGGCGGCGG 1077
QY 1081 GCGGAGGCGGATGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140
Db 1078 GCTGAGGCAATGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1137
QY 1141 GGCGCGCGCGCGCAACGTCAGTGTCTTCAACTGCGGCGGCGGCGGCGGCGGCGGCGGCG 1200
Db 1138 GGCCAAAGAGAAATTTGTTAAATGCTTTAAATTTGTTGCGCAAGAGGACACATAGCC 1197
QY 1201 TGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1260
```

Db 1198 TGCAGGGCCCTAGAAAAAAGGGCTGTGGAAGTGTGGAAGAGAGGACACCAATGAAG 1257
QY 1261 GACTGCACCGAGCGCCAGGCAACTTCCTGGGCAAGATCTGGCCAGCCACAAAGGCGGC 1320
Db 1258 GACTGCACCTAGAGACAGGCTAATTTTATAGGAAATTTGGCCTTCAACAAGGAGG 1317
QY 1321 CCGGGCAACTTCTGCGAAGCGGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1360
Db 1318 CCGGGAACCTTCATCCAGACAG-----GCCAGAGCGGTGAGCGCGCG 1359
QY 1381 CCGCGCGAGAGCTTCGCTTGGAGGAGACACCCCGCGCGCGCGCGCGCGCGCGCGCGCG 1440
Db 1360 CCAGCAGAGAGCTTCAGGTTGGGGGAGGAGACACCCCATCTCCGAGCCAGGAGGAGAA 1419
QY 1441 CCGGAGCCCTACCGGAGCGCGCTGACCGCGCTGCGCAGGCTGTTGGGAGCGCGCGCG 1497
Db 1420 GACGAGGAGCTGATCCCTCCCTAGCTTCCCTCAATCACTCTTTGGCAACGACCGCG 1476

RESULT 13

US-08-418-848A-9

; Sequence 9, Application US/08418848A

; Patent No. 5847096

; GENERAL INFORMATION:

; APPLICANT: SCHUBERT, MANFRED, HARMISON II,

; APPLICANT: GEORGE G., CHANG-JIE, CHEN, BANJERJEA, AKHIL

; TITLE OF INVENTION: DEFECTIVE, INTERFERING

; TITLE OF INVENTION: HIV PARTICLES

; NUMBER OF SEQUENCES: 77

; CORRESPONDENCE ADDRESS:

; ADDRESS: MORGAN & FINNEGAN, L.L.P.

; STREET: 345 PARK AVENUE

; CITY: NEW YORK

; COUNTRY: U.S.A.

; ZIP: 10154

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM PC COMPATIBLE

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: WORD PERFECT 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/418,848A

; FILING DATE: 07-APR-1995

; CLASSIFICATION: 526

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/936,849

; FILING DATE: 28-AUG-1992

; CLASSIFICATION: 526

; ATTORNEY/AGENT INFORMATION:

; NAME: RICHARD W. BORK

; REGISTRATION NUMBER: 36,459

; REFERENCE/DOCKET NUMBER: 2026-4091US2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-758-4800

; TELEFAX: 212-751-6849

; TELEX: 421792

; INFORMATION FOR SEQ ID NO: 9:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 7399 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; US-08-418-848A-9

Query Match

Best Local Similarity 42.6%; Score 642.2; DB 2; Length 7399;

Matches 998; Conservative 0; Mismatches 493; Indels 30; Gaps 3;

QY 1 ATGGCG 60

Db 751 ATGGGIGCGAGAGCGCTCGGTATTAGCGGGGAGAGATTAGATAAATGGGAAAAAATTCGG 810

QY 61 CTGGCG 120
Db 811 TTAAGGCCAGGGGGAAGAAACAATAATAAACTAAACATATAGTATGGCAAGCAGGGAG 870
QY 121 CTGGAGGCTTCGCGCTGAACCCCGCGCTGCTGGAGACCGCGCGAGGCTGCAAGCAGATC 180
Db 871 CTAGAAGGATTGCGAGTTAATCCTGGCTTTIAGAGACATCAGAAGCTGTAGACATA 930
QY 181 ATGAAGCAGCTGCAGCCCGCGCTGCAGACCGGCACCGAGGAGCTGCCGAGCCIGTACAAC 240
Db 931 CTGGGACAGCTACAACCATCCCTTCAGACAGGATCAGAAGAACTTAGATCATTTATATAAT 990
QY 241 ACCGTGGCCACCCCTGTACTCGCTGCACCGCGCGCATCGAGGTCCGCGACACCAAGGAGGCC 300
Db 991 ACAATAGCAGTCTCTATTGTGTGTCATCAAAAGGATAGATGTAAAAGACACCAAGGAAGCC 1050
QY 301 CTGGACAAGATCGAGGAGGAGCAGAAACAAGTCCAGCAGAAGACCCAGCAGGC----- 353
Db 1051 TTAGATAAGATAGAGGAAGAGCAAAACAAAAGTAAGAAAAAGGCACAGCAAGCAGAGCT 1110
QY 354 --CAAGGAGCGCGACGGCAAGGTGAGCCAGAAGTACCCCATCTGTGAGAAACCTGCAGGGC 411
Db 1111 GACACAGGAACAACAGCCAGGTGAGCCAAAATTTACCTATAGTGCAGAAACCTCCAGGG 1170
QY 412 CAGATGTTGCACCGAGGCGCATCAGCCCGCGCACCTGAACCGCTGGTGAAGGTGATCGAG 471
Db 1171 CAAATGTTACATCAGGCCATATCACCTAGAACITTTAAATGCATGGGTAAAGTAGTAGAA 1230
QY 472 GAGAAGGCTTCAGCCCGAGGTGATCCCGCATGTTACCGCGCTGAGCGAGGGCGCCACC 531
Db 1231 GAGAAGGCTTCAGCCCGAGGTGATCCCGCATGTTTCAGCATTTATCAGAGGAGGCCACC 1290
QY 532 CCCCAGGAGCTGAACACAGTGTGTAACACCGTGGCGCGCACCCAGCCCGCGCATGCGAGATG 591
Db 1291 CCACAAGATTTAAATACCATGCTAAACACAGTGGGGGACATCAACAGCCATGCAATG 1350
QY 592 CTGAAGGACACCATCAACGAGGAGGCGCGCGAGTGGGACCGCTGCACCCCGTGCAGGGCC 651
Db 1351 TTAAGAGAGACCATCAATGAGGAAGCTGCAGATGGGATAGATTGGATCCAGTGCATGCA 1410
QY 652 GGCCCGGTGCGCGCGCGCGAGATGCGCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCC 711
Db 1411 GGGCTATTGCACCGCGCGCGAGATGAGAGAACCAGGGGAAGTGACATAGCAGGAACACT 1470
QY 712 AGCACCTGCAGGAGCAGATGCGCTGGATGACCGACCGACCGCGCGCGCGCGCGCGCGGAC 771
Db 1471 AGTACCTTCAGGAACAATAAGTGGATGGATGACACATAATCCACCTATCCCGAGTAGGAA 1530
QY 772 ATCTAAGCGGTGGATCATCTGGCGCTGAACAAGATGCTGGGATGTACAGCCCGTG 831
Db 1531 ATCTATAAAGATGGATAATCTGGGATTAATAAATAAGTAGTAAGATGTATAGCCCTAAC 1590
QY 832 AGCATCTTGACATCGCTTC 891
Db 1591 AGCATCTTGACATAGACAAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGATC 1650
QY 892 TTCAAGACCTCG 951
Db 1651 TATAAAACICTAAGAGCG 1710
QY 952 CTGCTGTGAGAAACCGCAACCG 1011
Db 1711 TTGTTGGTCCAAAATGCGAACCAGATTGTAAGACTATTTTAAAGCATTTGGGACAGGA 1770
QY 1012 GCCACCTGGAGGAGATGATGACCGCGCTGCCAGGCGCTGGCGCGCGCGCGCGCGCGCGCG 1071
Db 1771 GCGACACTAGAAGAAATGATGACAGCATGTGCGAGGTGGGGGAGCGCGCGCGCGCGCGCG 1830
QY 1072 CGCGTGTGCGCGGAGGATGAGCGCGG--CCAACAGCGTGAACATCATGATGAGAGAAG 1128
Db 1831 AGAGTTTGGCTGAAGCATGAGCGCAAGTAACAAATCCAGCTACCATATGATACAGAAA 1890
QY 1129 AGCAACTTCAAGGGCG 1188

Db 1891 GGCATTTTAGAACCAAGAAAGACTGTTAAGTGTTCATATGTCACAGAGAGCCGAC 1950
Qy 1189 ATGCCAAGAAGTGGCGGCGCCGCGGCAAGAAAGGGGTGCTGSAAGTGGGCAAGGAGGSC 1248
Db 1951 ATAGCCAAAATGTCAGGGCCCTAGGAAAAGGGCTGTTGANAATGTGGAAGGAAGGA 2010
Qy 1249 CACCAGATGAAGGACTGCACCGAGCGCCAGGCAATTCCTGGGCAAGATCTGGCCAGC 1308
Db 2011 CACCAATGAAGATTGTCCTGAGAGACAGGCTAATTTTACGGAAGATCTGGCTTCC 2070
Qy 1309 CACAAGGGCGGCGCCGCGCAACTCTCTGCAGAACCGGAGCGAGCGCGCGCCGCGCCAGCGTG 1368
Db 2071 CACAAGGGAAGCGCAGGGAATTTCTTCAGAGCAG-----ACCAAG 2112
Qy 1369 CCACCGCGCCCGCGCGGAGAGCTTCGGTTCGAGGAGACCAACCGCGCGCGCAAGCAG 1428
Db 2113 CCAACAGCGCCCGCAGAGAGAGCTTCAGGTTTGGGGAAGAGACAAACCTCTCAG 2172
Qy 1429 GAGCCCAAGGACCGGAGCGCTACCGCGAGCGCCCTGACCGGCGCTGCGCAGGCTGTTCGGC 1488
Db 2173 AAGCAGGAGCGGATAGACAAGGAAGTGTATCTTAGCTTCCTCAGATCATCTTTGGC 2232
Qy 1489 AGCGGCGCCCTGAGCCAGTAA 1509
Db 2233 AGCGACCCCTCGTCACAATAA 2253

RESULT 14

US-08-188-583-5
; Sequence 5, Application US/06188583
; Patent No. 5851813
; GENERAL INFORMATION:
; APPLICANT: Desrosiers, Ronald C.
; TITLE OF INVENTION: PRIMATE LENTIVIRUS VACCINES
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 507 or 55SX
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
; SOFTWARE: WordPerfect (Version 5.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/188,583
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/727,494
; FILING DATE: July 9, 1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/551,945
; FILING DATE: July 12, 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: Reg. No. 5851813 29,366
; REFERENCE/DOCKET NUMBER: 00246/079002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9709
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

Query Match 42.6%; Score 642.2; DB 2; Length 9709;
Best Local Similarity 65.6%; Pred. No. 1.9e-86;
Matches 998; Conservative 0; Mismatches 493; Indels 30; Gaps 3;
Qy 1 ATGGGCGCCCGCGCCAGCATCTCTGCGGGCGAGAGCTGGACAACTGGGAGAGATCCGC 50
Db 790 ATGGGTGGAGAGCGTGGTATTAAAGCGGGGGAGAAATTAGATAAATGGGAAAAAATTCGG 849
Qy 61 CTGCGCCCGCGCGGCAAGAGCACTACATGCTGAGACACCTGGTGGGCGCCAGCGAG 120
Db 850 TTAAGGCGAGGGGAAAGAAACAATAAACTAAACATATAGTATGGGCAAGCAGGGAG 909
Qy 121 CTGGAGGCTTCGCCCTGAACCCCGGCTGCTGGACACCGCGGAGGCTGCAAGCAGATC 180
Db 910 CTAGAACGATTGCGAGTTAATCTTGGCTTTTAGAGACATCAGAAAGGCTGTAGACAATA 969
Qy 181 ATGAAGCAGCTGCAGCGCCCGCTGCGAGACCGCGGAGGAGCTGCGCAGCCCTGTACAAC 240
Db 970 CTGGGACAGCTACACCATCTCTTICAGACAGGATCAAGAAGACTTAGATCATTTATATAAT 1029
Qy 241 ACCGTGGCCACCTGTACTGCTGGTGGCGCGGCTGCGGCTGCGGAGGCTGCGGAGGAGGCG 300
Db 1030 ACAATAGCAGTCTCTATTGTGTCATCAAGGATAGATGTAAGACACCAAGGAGGCC 1089
Qy 301 CTGGACAGATCGAGGAGGAGCAGACAAGTCCAGCAGAGAGACCAAGCAGCAGCC----- 353
Db 1090 TTAGATAAGATAGAGGAGAGCAAAACAAGTAAGAAAGGCGACAGCAAGCAGCAGCT 1149
Qy 354 --CAAGGAGCGCGACGCAAGGAGGAGCAAGACTACCCCATCTGTCAGANCCISGAGGGC 411
Db 1150 GACACAGGAACAACAGCCAGGTCAGCAAAATACCTTAAGTGGAGAGACCTCCAGGGG 1209
Qy 412 CAGATGGTGCACACAGGATGTTGACACCGTGGCGGCGCCAGCAGCGCCCATGCAAGATG 471
Db 1210 CAAATGGTACATCAGGCCATATCACCTAGAATTTAAATGCTGGTAAAGTAGTAGAA 1269
Qy 472 GAGAAGSCCTTCAGCCCGGAGGTGATCCCCCATGTTTACCCGCTGAGCGAGGCGGCCACC 531
Db 1270 GAGAAGSCCTTCAGCCCGAGAGTAATACCCATGTTTTCAGCATTTATCAGAGGAGGCCACC 1329
Qy 532 CCCCAGGACCTGAACACGATGTTGACACCGTGGCGGCGCCAGCAGCGCCCATGCAAGATG 591
Db 1330 CCACAGATTTAAATACCATGCTAAACACAGTGGGGGAGACATCAAGCAGCCATGCAATG 1389
Qy 592 CTGAAGGACACCATCAACGAGGAGGCGCGGAGTGGGACCGCTGCACCCCGTGCAGGGCC 651
Db 1390 TTAAGAGAGACCATCAATGAGGAAGCTGCAGATGGGATAGATTGTCATCCAGTGCATGCA 1449
Qy 652 GGCCCGCTGGCGCCAGATGCGGAGACCCCGCGGAGCGGAGCGGAGCGCGCGGCGCAC 711
Db 1450 GGGCCTATTGCACCGGCGAGATGAGAGAACCAAGGGAAGTGACATAGCAGGAAGTACT 1509
Qy 712 AGCACCCTGCAGGAGCAGATCGCCCTGATGACCAAGCAACCCCGCTGCGCGGAG 771
Db 1510 AGTACCCCTCAGGAACAATAAGGATGGATGACACATAATCCACCTATCCAGTAGGAGAA 1569
Qy 772 ATCTAAGCGGTGGATCATCTCTGGCTGAACAAGATGTCGGGATGTACAGCCCGTG 831
Db 1570 ATCTATAAAGATGGATAATCTCTGGGATTAATAAATAGTAAGAATGTATAGCCCTACC 1629
Qy 832 AGCATCTGGACATCCCGCAGGCGCCCAAGGAGCCCTTCGCGAGTACGTGGACCGCTTC 891
Db 1630 AGCATCTGGACATTAAGACAAGGAGCAAGGAGGAGGAGCCCTTTAGAGACTATGTAGACCGATTC 1689
Qy 892 TTCAGACCCCTGCGCGGCGGAGCAGGCGCCAGCAGCGGAGCTGAAGAAGTGCATGACCGAGCC 951
Db 1690 TATAAAACICTAAGAGCGGAGCAAGCTTCACAAGAGAGTAAAAAATTTGGATGACAGAAACC 1749
Qy 952 CTGCTGGTGCAGAACGCGCAACCCCGACTGCAAGAGCAATCTGCGGCTCTCGGCCCGGC 1011
Db 1750 TTGTTGGTCCAAAATGGCAACCCAGATGTGAAGACTATTTTAAAGCATTTGGGACCAGGA 1809

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2003, 17:31:08 ; Search time 3435.15 Seconds
(without alignments)
10676.533 Million cell updates/sec

Title: US-09-475-704A-4
Perfect score: 1509
Sequence: 1 atgggcqccgcgcagcat.....gcggcccccctgaqccagta 1509

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :		EST:	
1:	em_estba:	*	
2:	em_esthum:	*	
3:	em_estin:	*	
4:	em_estmu:	*	
5:	em_estov:	*	
6:	em_estpl:	*	
7:	em_estro:	*	
8:	em_htc:	*	
9:	gb_est1:	*	
10:	gb_est2:	*	
11:	gb_htc:	*	
12:	gb_est3:	*	
13:	gb_est4:	*	
14:	gb_est5:	*	
15:	em_estfun:	*	
16:	em_estom:	*	
17:	em_gss_hum:	*	
18:	em_gss_inv:	*	
19:	em_gss_pln:	*	
20:	em_gss_vrt:	*	
21:	em_gss_fun:	*	
22:	em_gss_mam:	*	
23:	em_gss_mus:	*	
24:	em_gss_pro:	*	
25:	em_gss_rod:	*	
26:	em_gss_phg:	*	
27:	em_gss_vrl:	*	
28:	gb_gss1:	*	
29:	gb_gss2:	*	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	83.2	5.5	933	29	CC391778
2	81.6	5.4	1219	14	CD500131
3	81	5.4	749	13	BQ744212
4	80.6	5.3	1674	10	BE636712

C	5	79.8	5.3	1201	13	BX356664	BX356664
	6	79.8	5.3	2299	11	AY106831	AY106831
	7	76.4	5.1	759	10	BF259495	BF259495
	8	76.4	5.1	1308	11	AY104577	AY104577
	9	75	5.0	509	10	BG314500	BG314500
C	10	74.4	4.9	1170	14	CD496623	CD496623
	11	74.4	4.9	2598	11	AY103647	AY103647
	12	74.2	4.9	649	14	CB870888	CB870888
	13	74	4.9	637	12	B1960118	B1960118
	14	73.8	4.9	562	10	BF484304	BF484304
C	15	73.2	4.9	694	12	BJ261281	BJ261281
	16	72.6	4.8	575	13	BQ842001	BQ842001
	17	72.6	4.8	501	12	BJ290713	BJ290713
	18	72.6	4.8	652	12	BJ292215	BJ292215
C	19	72.6	4.8	700	12	BJ295604	BJ295604
	20	72.6	4.8	712	13	BQ842240	BQ842240
C	21	72.6	4.8	718	12	BJ297319	BJ297319
C	22	72.6	4.8	753	12	BJ298801	BJ298801
	23	72	4.8	731	13	BQ752847	BQ752847
	24	71.8	4.8	1491	11	AY104431	AY104431
C	25	71.6	4.7	655	12	BJ300074	BJ300074
	26	70.8	4.7	1327	10	BE636532	BE636532
	27	70.6	4.7	671	12	BI718023	BI718023
	28	70.6	4.7	736	13	BQ295509	BQ295509
	29	70.4	4.7	764	13	BQ804819	BQ804819
C	30	70.2	4.7	648	10	BE517305	BE517305
C	31	69.6	4.6	690	14	CD495520	CD495520
C	32	69.4	4.6	709	12	BJ283977	BJ283977
	33	69	4.6	640	29	BZ643623	BZ643623
C	34	68.6	4.5	744	29	CC337622	CC337622
C	35	68.4	4.5	694	12	BJ265176	BJ265176
	36	68.4	4.5	841	29	CC415062	CC415062
C	37	68.4	4.5	848	29	CC415063	CC415063
	38	68.2	4.5	509	12	BJ549094	BJ549094
	39	68.2	4.5	617	13	BQ739819	BQ739819
	40	68.2	4.5	642	12	BM575079	BM575079
	41	68.2	4.5	664	13	BQ762267	BQ762267
	42	68	4.5	709	13	BQ837986	BQ837986
C	43	67.8	4.5	638	12	BJ254941	BJ254941
C	44	67.8	4.5	657	12	BQ215186	BQ215186
	45	67.8	4.5	763	13	BQ100470	BQ100470

ALIGNMENTS

RESULT 1	CC391778	CC391778	933 bp	DNA	linear	GSS 19-MAY-2003					
LOCUS	PUHIJ70TB	ZM_0.6_1.0_KB	Zea mays	genomic clone	ZMMBtra458K20,						
DEFINITION	genomic survey sequence.										
ACCESSION	CC391778										
VERSION	CC391778.1	GI:30871868									
KEYWORDS	GSS.										
SOURCE	Zea mays										
ORGANISM	Zea mays										
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.										
AUTHORS	1 (bases 1 to 933) Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and Bennetzen,J.										
TITLE	Maize Genomics Consortium										
JOURNAL	Unpublished										
COMMENT	Other_GSSs: PUHIJ70TD Contact: Cathy Whitelaw TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA Tel: 301-838-5843 Fax: 301-838-0208 Email: whitelaw@tigr.org Seq primer: TR Class: sheared ends.										

FEATURES
Source Location/Qualifiers
1..933
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="P73"
/db_xref="taxon:4577"
/clone_lib="ZM0.6.1.0_KB"
/note="vector: PCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
Cot selected genomic DNA library"
BASE COUNT 198 a 276 c 309 g 150 t
ORIGIN
Query Match 5.5%; Score 83.2; DB 29; Length 933;
Best Local Similarity 45.7%; Pred. No. 0.00019;
Matches 329; Conservative 0; Mismatches 388; Indels 3; Gaps 1;
QY 189 GCTGCAGCCGCCCTGCAGACCGGGACCGGAGGCTGCCAGCCTGTACAAACACCGTGGC 248
DB 17 GCCGAGGGCGGGTGTCCGGCACCGAGAGGGCGCGTGCAGCAAGCAGGAGCGTGTGGT 76
QY 249 CACCCTGTACTGCTGCACGCGCGGATCGAGGTCCGGGACACCAAGGAGGCCCTGGACAA 308
DB 77 CAGGTGCACAGCACCCAGAGGATTCGACGACGCGGCTCAAGGCCGCCAAGAACCGGCTGGT 136
QY 309 GATCGAGGAGGAGCAGAACAAAGTCCACGACGAGACCCAGCAGGCGCAAGGAGGCGGACGG 368
DB 137 GGTGTGAGTTCCGGCGGAGCGACAGCGAGGACAGCAGCAGATATACCCGACCATGGT 196
QY 369 CAAGGTGAGCCAGAACTACCCCATCGTCGAGAACCTGCAGGGGCCAGATGGTGCACGAGGC 428
DB 197 GCAGCTCAGCCGCACCTCCGGCGGACGTGGACTTCTCTGCTGCTCTGCTGCTGCTGCTG 256
QY 429 CATCAGCCCGCCACCTGACCGCTGGGTTGAGGTGATCGAGGAGAGGCGCTTACGCC 488
DB 257 GGAGGCCACCGAGGAGGCTGTTCGGCGGAGGAGGCGTACGAGAGGTGCCCCACTTCAACT 316
QY 489 CGAGGTGATCCCATGTTACCCGCTGACCGGCTGACGAGGAGGCGGCCACCCCGCAGGACCT 545
DB 317 CTACAAGGCGCGCGAGAGGTGCACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 376
QY 546 CACGATGTTCAACACCGTTGGCGCGGCGGACGAGGCGGCGGATGCTGAGAGGAGGAGAT 605
DB 377 CGAGCTCTCTACTACGGGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 436
QY 606 CAACGAGGAGGCGCGGAGTGGGACCGCTGCACCGGCTGACGAGGAGGAGGAGGAGGAGG 665
DB 437 CGTGGAGGCGCTCATCGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 496
QY 666 CGGCCAGATGGCGGACCGCGGCGGCGGACATCGCGGCGGCGGCGGCGGCGGCGGCGGAG 725
DB 497 GGGCTCAGGACACTCGGGCGGCTCGGTCAGGTGTACCGCCACCGTGTGAGGCTGCGCG 558
QY 726 GCAGATCGCTGGATGACGAGCAACCGCGCGTGCACCGGCTGCGGCGGCGGCGGCGGCG 783
DB 557 GTCCATGTCGACAAACCGCTCTTCCGCGGCGATGAGCGGCGGAGGAGGAGGAGGAGG 616
QY 786 GATCATCTGGGCGCTGACAAAGATCGTGTGGATGTACAGCCCGCGTGAACATCCTGACAT 843
DB 517 GGAGTCTCAGGGCCATGAAGATCGTGTGAGGTGCCCACTTCTGCTTCATCAGGAGCGG 676
QY 846 CCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACCC13CG 905
DB 677 CCAGATCGTCGGCCGCTACGTCGGCTCCGGCAAGGGGAGGCTGCTGCTGAGATCCTCCG 736

RESULT 2
CD500131/c
LOCUS
DEFINITION
CD43-C08.yld-s SHGC-CDA Gasterosteus aculeatus cDNA clone
ACCESSION
CD500131
VERSION
CD500131.1 GI:31427162

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
EST.
Gasterosteus aculeatus (three spined stickleback)
Gasterosteus aculeatus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes;
Gasterosteidae; Gasterosteus.
1 (bases 1 to 1219)
Kingsley,D.M., Peichel,C., Balabhadra,S., Grimwood,J., Dickson,M.,
Schmutz,J. and Myers,R.M.
Expressed sequence tags from Gasterosteus aculeatus
Unpublished
Contact: Kingsley, DM
HHMI and Department of Developmental Biology
Stanford University School of Medicine
Beckman Center B300, 279 Campus Drive, Stanford, CA 94305-5329, USA
Tel: 650 725 5954
Fax: 650 725 7739
Email: kingsley@cngm.stanford.edu
Plate: 43
High quality sequence start: 4
High quality sequence stop: 792.
Location/Qualifiers
1..1219
/organism="Gasterosteus aculeatus"
/mol_type="mRNA"
/strain="Salinas river, CA"
/db_xref="taxon:69293"
/clone="CDA43-C08"
/sex="mixed male and female"
/tissue_type="heads and internal organs combined"
/dev_stage="adult"
/clone_lib="SHGC-CDA"
/note="vector: lambda ZAP Express/pBK-CMV; Site_1: EcoRI
(5' adaptor); Site_2: XhoI (3' linker primer); The mixed
organ cDNA library was generated using the ZAP-cDNA method
by Stratagene. First strand cDNA synthesis was primed with
a 50 bp linker primer containing an oligo dT sequence
preceded by a synthetic XhoI site. 5 prime adaptors were
used containing an EcoRI cohesive end. The finished cDNAs
were inserted in to the ZAP express vector
unidirectionally in the sense orientation with respect to
the lacZ promoter of pBK-CMV. An amplified library was
prepared from approximately 3 million primary clones in
the lambda ZAP Express vector. In vivo excision was then
used to generate individual pBK-CMV phagemid clones for
EST sequencing."
BASE COUNT 201 a 440 c 390 g 188 t
ORIGIN
Query Match 5.4%; Score 81.6; DB 14; Length 1219;
Best Local Similarity 47.7%; Pred. No. 0.00036;
Matches 373; Conservative 0; Mismatches 394; Indels 15; Gaps 4;
QY 616 CGCCGAGTGGGACCGCTGCACCCCGTGCAGGCGGCGCGCGCGCGCGCGCGCGCGCG 677
DB 917 CGGCTCGTGGCGGACTACGGCGGGGTACGGCGGACGCCCTGCTCAACCGGCTCGCTGGG 858
QY 678 CGACCCCGCGCGGCGGAGC-----ATGCGCGGCGCGCGCGCGCGCGCGCGCGCG 731
DB 857 CGACTACGGCGGCGGCGGAGCGCCCTGCTCAACCGGCTGCTGGCGGCGGCGGCGGCG 798
QY 732 CGCCTGGATGACGACACCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 791
DB 797 CGCCCTGCTCAACGGCTCGCTGGCGGACTACGGCTACGGCGGCGCGCGCGCGCGCGCTC 738
QY 792 CTTGGGCTGACAAAGATCGTGGGATGTACAGCCCGCGCGCGCGCGCGCGCGCGCGCG 851
DB 737 GCTGGCGGACTACGGCGGAGCGCCCTGCTCAACGGCGGCGGCGGCGGCGGCGGCGGCG 678
QY 852 GGGC---CCCAAGGAGCCCTTCGCGGACTACGTGGGAGCGCTTCTTCAAGACCCCTGC---G 905
DB 577 CGCCCTGCTCAACGGCTCGCTGGCGGACTACGGCGGCGGCGGCGGCGGCGGCGGCGCTCA 618

BE536712
LOCUS BE536712 1674 bp mRNA linear EST 03-JAN-2002
DEFINITION rockefeller.0.42 Mastigamoeba balamuthi lambda ZAP 11 library
Mastigamoeba balamuthi cDNA similar to vacuolar ATP synthase
subunit B (EC 3.6.1.3), mRNA sequence.
ACCESSION BE536712 GI:9919823
VERSION BE536712.1
KEYWORDS EST.
SOURCE Mastigamoeba balamuthi
ORGANISM Mastigamoeba balamuthi
Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.
REFERENCE 1 (bases 1 to 1674)
AUTHORS Baptiste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,
Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and
Philippe,H.
TITLE The analysis of 100 genes supports the grouping of three highly
divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
MEDLINE 21819461
PUBMED 11830664
COMMENT Contact: Muller Miklos
Laboratory of Biochemical Parasitology
The Rockefeller University
1230 York Avenue, New York, NY 10021, USA
Email: mmuller@rockvax.rockefeller.edu
Insert Length: 1674 Std Error: 0.00
POLYA-Yes.
FEATURES
Location/Qualifiers
1..1674
/organism="Mastigamoeba balamuthi"
/mol_type="mRNA"
/strain="ATCC 30984"
/db_xref="taxon:108607"
/clone_lib="Mastigamoeba balamuthi lambda ZAP 11 Library"
/note="syn: Phreatamoeba balamuthi"
BASE COUNT 348 a 561 c 493 g 269 t 3 others
ORIGIN
Query Match 5.3%; Score 80.6; DB 10; Length 1674;
Best Local Similarity 44.1%; Pred. No. 0.00054;
Matches 615; Conservative 0; Mismatches 756; Indels 22; Gaps 6;
QY 59 GCCTGCGCCCGCGGCAAGAGCACTACATGCTGAGCACTGCTGTGGGCCAGCGCG 118
DB 136 GCGCTGCTGGTCTGCTGACACATCAAGAGCGCGAGTTCGACAGATGCTCAAGCTGA 195
QY 119 AGCTGAGGGGCTTCGCGCTGAACCGCGGCTGCTGGAGACCGGAGGGTGCAGCAGA 178
DB 196 CGCTGGCGGACGGGAGCGGTGGCGCGGCGAGTGTCTGGAGTGGCGCGGACAGCGCG 255
QY 179 TCATGAAGCAGCTGCAGCGCGCGCTGCAGACCGGACCGAGAGCTGCGCGGCTGTACA 238
DB 256 TCGTGCAGGTCCTTGAGGGCAGCTCGGGCATCGACGCCAAGACACGCACTCGGAGTCA 315
QY 239 ACACGGTGGCCACCTGTACTGCTGCAGCGCGCGCATCGAGTCCGGACACCAAGGA-- 296
DB 316 CGGCGCACATCATCGGCATCCCGTCAAGGAGGAGACATGCTGGCGCGGTCTTCAAGGCT 375
QY 297 -GGGCTGGACAGATCGAGGAGGAGCAGACAGTCCGACAGAGACCGACAGCGGCA 355
DB 376 CGGCGCGCGCGTGCAGACAGGCGCGGCGGTCTGCGCGAGGAGTTCCTCGACATCGAGG 435
QY 356 AGGAGCGCGGCGCAAGGTGAGCGCAGAACTACCGCATCTGTGAGAACCTGCGAGGGCGCAGA 415
DB 436 GCCAGCCCATCAACCGGTACAGCGGTGTGTACCGGAGGAGATGATCCAGAGCGGCATCA 495
QY 416 TGGTGCACCAAGGCCCATCAGCCCGCGCACCTGACCGCTGGGTGAAGGATCGAGGAGA 475
DB 496 GCGCCATCGACACATGAAGTCACTGATCGCCCGTGGCCAGAGATCCCGCTGTTCAGCGCG 555
QY 476 AGGCTTTCAGCCCGGAGGTGATCCCCAIGTTTACCGCGCTGACGAGGGCGCCACCCGCC 535
DB 556 CCGGTCTGCCGCACACAGAGATGTCCCGCGCAGATCTGCAGGCGAGCGCGCTGGTGGCGCC 615

QY 536 AGGACCTGAACACGATGTTGAACACCGTGGCGGGCCACCAGGCGGCCCAIGCAGATGCTGA 595
DB 616 TCCAGAGCAGCAGAGAGGTGCAGGACGACCACGAGGACAACTTCGCCATCTGTTCGGCG 675
QY 596 AGGACACCAATCAACGAGGAGGCGCGCGAGTGGGACCGGCTGCACCCCGTGCAGGCGCG 655
DB 676 CCATGGGTGTCAACATGGAGACTGCCAGGTTCTTCGGCCAGGACTTCGAGGAGAACGGCT 735
QY 656 CCGTGGCCCCCGGCCAGATGCGGACCCCGCGGCGGCGAGGACATCGCGGCGGCCACGACGA 715
DB 736 CGATGGAGCGGTGCACGCTGTTCTCAACCTCGCCCAACCCACCGATCGAGCGCATCA 795
QY 716 CCCTGCAGGAGCAGATCGCTGGATGACGACCAACCCCGCTGCCGTGGCGGACATCT 775
DB 796 TCACGCGCGGTCTG---GGCTCAGCAGCGGCGGAGTACCTCGGCTACACGTCGAGATGC 852
QY 776 ACAGCGGTGGATCATCTCTGGCT---GAACAAGATCTGTGGGATGTACAGCCCG 829
DB 853 ACGTGTGTCTATCTCTGACGGACATGTCTGTACGCGGATGCCCTGCGTGAGNTCTCGG 912
QY 830 TGACATCTTGGACATCCGCCAGGCGCCCAAGAGCGCTTCCCGGACTACGTGGACCGCT 889
DB 913 CGGCGCGTGGAGGTCCTTGGCGTCTGTGTTACCCCGGTATCATGTACACGGATCTCG 972
QY 890 TCTTCAAGACCTTGGCGCGGAGGAGGCGCACCGAGGACGTGAAGAACTGGATGACCGAGA 949
DB 973 CGACCATCTACGAGCGTGGCGGCGTG---TCAGGGCAGGAACGGTCTGATCAGCGAGC 1029
QY 950 CCGTGTGTGTCAGAACGCCAACCCCGGACGACGACGACCATCTCTCGCGCTCTCGGCCCG 1009
DB 1030 TCCCATCTCTACGATGCCCAACGACGACATCACCCACCCCTCCCTGACCTCAGCGGT 1089
QY 1010 GCGGCACTTGGAGGAGATGATGACCGCTGCGAGGCGTGGCGGCGCGCCGCGCACAAAG 1069
DB 1090 ACATCACCAGGCGCAGATCTACGTCGACAGGAGCTGCACACAGGCGAGATCTACCCGC 1149
QY 1070 C---CGCGGTGTGCGGAGGCGATGAGCCAGGCGCAACAGCGTGAACATCATGATGCAGA 1126
DB 1150 CCATCAACGTCGTGCGCTGCTCAGCGCTCTGATGAAGAGCGGCATCGCGAGGCGCATGA 1209
QY 1127 AGAGCAACTTCAAGGCGCCCGCGGCGCAACGTCAGTGCCTTCAACTGCGGCAAGGCGCC 1186
DB 1210 CGGCGGACGACCCACTCGGAGGTGTGAGACCGAGCTGTACGCCCACTACGCCATCGGCNAGG 1269
QY 1187 ACATCCCAAGAACTGCCGCGCGCGCGCGCAAGAGGCGCTGTGGAAGTCCGCAAGGAGG 1246
DB 1270 ACAGATGGCCATGAAGGCTGTCTCGGTGAGGAGGCGCTGTGCGCCCGAGGACAGCTCG 1329
QY 1247 GCCACAGATGAAGGACTGCACCGAGCGCCAGGCGCAACTTCTTGGGCAAGATCTTGGCCCA 1306
DB 1330 CCGTGGAGTTCCTCGAGAAGTTCGAGCAGAAAGTTCATCACCGAGGGCTACACAGAGTCCG 1389
QY 1307 GCCAGAGGCGCGCGCGCAACTTCCCTGACAGACCGCGAGGCGCGCGCGCCCGCCNCCG 1366
DB 1390 GCGACATCTACAGTCCGCTCGACATCGCCTGGAGCGCTCTCGG---CCAGTTCGCCAAGG 1445
QY 1367 TGCCACCGCGCGCGCGCGCGGAGAGCTTCGGCTTCGAGGAGACACCCCGCGCCCGCAAGC 1426
DB 1446 AGCTCTCAGCGCTATCAGCGGAGAGAGACCCCTCAACANGTACTACAGCAGGAGCGCAAGN 1505
QY 1427 AGGACCCCAAGGA 1439
DB 1506 AGGAGCACCAAGTA 1518

RESULT 5
BX356664/c
LOCUS BX356664 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0DI015YB03 3-PRIME, mRNA sequence.
ACCESSION BX356664
VERSION BX356664.1 GI:30378083

KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
AUTHORS Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@litech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DI015CA02NP1.
FEATURES Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI015B03"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dI)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 116 a 88 c 93 g 398 t 506 others
ORIGIN
Query Match 5.3%; Score 79.8; DB 13; Length 1201;
Best Local Similarity 9.6%; Pred. No. 0.00071;
Matches 59; Conservative 303; Mismatches 254; Indels 0; Gaps 0;
QY 894 CAAGACCCTGGCGCCGAGGAGCCAGCCAGGACCTGGAAGAACTGGATGACCGAGACCC 953
DB 1201 VAAVSSSAVSSSSVAAASSVAAASSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 1142
QY 954 GCTGCTGCAGAACGCCAACCCCGACTGCAAGACATCTCTGGCGCTCTCGGCCCGGCG 1013
DB 1141 SVVSSSAAASSSSSSSAAAVVAAASSVSVATSSSAAAAAASSSSSSVSSAAAA 1082
QY 1014 CACCCTGGAGGAGATGATGACCGCTGCCAGGGCTGGCGGCGCTCTCGGCCCGGCG 1073
DB 1081 ASVSSSSVAAVSAVAAAVVSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 1022
QY 1074 CGTCTGGCCGAGCGGATGAGCCAGGCCAACAGCGGTGAACATCATGTGACAGAGAGCAA 1133
DB 1021 SSSANSVSSSVSSSAAASSVAAAVSSAAVSSABAAAAAASVAAASSSSSSSSSSSS 962
QY 1134 CTTCAAGGGCCCCGGCGCAACGTCAAGTCTTCAACTGCGGGAAGGAGGGCCACATCG 1193
DB 961 MSSAVSSSSSVSSSSSAVSSSSSSSAAASVSSSAAVSSSVAAASSSSSSSSSVAA 902
QY 1194 CAAGNACTGCGCGCCCGCCCGCCAGAAAGGGCTGCTGGAAGTGGGCAAGGAGGCCACCA 1253
DB 901 SAAAVASAVSAAAVVSWSSVVAASSASVASASVSSRSSSAAARASSAGSAAVAAWASS 842
QY 1254 GATGAAGGACTGCACCGAGCGCCAGGCCAACCTCTCGGCAAGATCTGGCCCGCCACAA 1313
DB 841 SSVAVVSAAVSSKASNAASSAAVSSSSSSSSSSSAGAVSSSAKSVASSAAVSSAGSS 782
QY 1314 GGGCGCCCGCGCAACTTCTCTGCAAGACCGCAGGAGCCCGCCCGCCCGCCCGCCAC 1373
DB 781 VSSVSVRSSSVARVAASSVAAAGSAVASAASSASVASSSVSMAAAGSAAVSVASSSAV 722
QY 1374 CGCCCCCCCCCGCAGAGCTTCCTCTCGAGGAGACCAACCCCGCCCGCCCGCCCGCAG 1433
DB 721 SAAAASSSVSAASASSSSAAASSSSSSAAAVVSSSSSASSSSSVARSVAVSSASS 662
QY 1434 CAAGACCGGAGGCGCTACCGGAGCCCTGACCGCCCTGCGCAGACCTGTTCGGCAGCGG 1493

DB 661 SASSSAAVAAASANSANNSSSSSSNNNNNAASSSSSANAANCNSSNNSSSSNN 602
QY 1494 CCCCCTGAGCCAGTAA 1509
DB 601 NSSNSNANNNNANAAA 586
RESULT 6
AY106831
LOCUS AY106831
DEFINITION Zea mays P00142079 mRNA sequence.
ACCESSION AY106831
VERSION AY106831.1 GI:21209909
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 2299)
REFERENCE Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
JOURNAL Unpublished (2002)
REFERENCE 2 (bases 1 to 2299)
AUTHORS Coe,E.H.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.
FEATURES Location/Qualifiers
1..2299
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:638374"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACS in conjunction with the Maize
Mapping Project"
BASE COUNT 475 a 790 c 658 g 376 t
ORIGIN
Query Match 5.3%; Score 79.8; DB 11; Length 2299;
Best Local Similarity 43.9%; Pred. No. 0.00077;
Matches 495; Conservative 0; Mismatches 617; Indels 15; Gaps 3;
QY 122 TGGAGGGCTTCGCCCTGACCCCTGACCCCGGCTGCTGGAGACCGCGGAGGGTCAAGCATCA 181
DB 767 TCGAGCGCGCGCCATCATCGAGCACATCTGGATGGCAGCTCTTCATGAAGCAGCCCA 826
QY 182 TGAAGCAGCTGCAGCCCGCCCTGCAGACCGGCCAGGAGAGCTGCGCAGCCTGTACAACA 241
DB 827 AGGAGGTGAACGCCCATGGACCCCGCTGCIGRAGCCGGAAGCAGGACAGGTACCGCTCCG 886
QY 242 CCGTGGCCACCTGTACTGCTGCACCGCGGCATCGAGGTCCGCGACACCAAGAGGCC 301
DB 887 CGTCGCCCGCAGTGGCTCGGCCCGCCAGATCGAGGTATCGCGCGCCGCGCAGTCCATCG 946
QY 302 TGGACAAGATCGAGGAGGAGCAGAACAGTCCCAGCAGAGACCCAGCAAGCCAGGAGG 361

Db	881	CCGCTGGAGCTGCTCAAGTTGCGCCATCGGCTACGGCGCTCTCGGGCAAGGCGCTGGAACA	940
QY	1245	GGGCCACCATGAAGGACTGCGACCGCCAGGCGCAACTTCCTGGGCAAGATCTGGCC	1304
Db	941	CATCAACAAGAAGACGGCCTTACCAACCGCACCGACTACGGCAAGGGCGAG-----	992
QY	1305	CAGCCACAAGGGCGCCGCCGCCCAACTTCCTGCAGAACCGCACCGCGCGCGCGCCGAC	1364
Db	993	-----CGAGAGCGCAGTGGGCCACGGCACAGAGAGCGCTGCACGGCTCAACCCAGGCCAC	1048
QY	1365	CGTCCCCACGCCCCCGCCCGAGAGCTTCGGTTGAGGAGAGCACACCCCGCCCGCAA	1424
Db	1049	CGCCACCTCCGACCTCTTCGGCGACAACCGACGGGCTACCGCGAGCTGTGGAGGCTCGCGGA	1106
QY	1425	GCAGGAGGCCAAGGACCGCGAG	1446
Db	1109	GCAGGCGGCCAAGCGCGCGAG	1130

RESULT 9
BG314500
LOCUS
DEFINITION
WHE2495_E11_I2L2S Triticum monococcum early reproductive apex cDNA library Triticum monococcum cDNA clone WHE2495_E11_I2, mRNA sequence.
ACCESSION
BG314500
VERSION
BG314500.1 GI:13116303
KEYWORDS
EST.
SOURCE
Triticum monococcum
ORGANISM
Triticum monococcum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poaceae : Triticeae; Triticum.
REFERENCE
1 (bases 1 to 509)
AUTHORS
Anderson, O.D., Chao, S., Dubcovsky, J., Echenique, V., Han, P.S., Hsia, C.C., Kang, Y., Lazo, G.R., Miller, R., Rausch, C.J., Seaton, C.L., Stamova, B. and Tong, J.C.
TITLE
The structure and function of the expressed portion of the wheat genomes - Early reproductive apex cDNA library from Triticum monococcum

JOURNAL
 COMMENT
 Unpublished
 Contact: Clin Anderson
 US Department of Agriculture, Agriculture Research Service, Pacific
 West Area, Western Regional Research Center
 800 Buchanan Street, Albany, CA 94710, USA
 Tel: 5105595773
 Fax: 5105595818
 Email: anderson@pw.usda.gov
 Sequence have been trimmed to remove vector sequence and low
 quality sequence with phred score less than 20
 Seq primer: Stratagene SK primer.

```

FEATURES
    source
        location/Qualifiers
            1. .509
                /organism="Triticum monococcum"
                /mol_type="mRNA"
                /cultivar="DV92"
                /db_xref="taxon:4568"
                /clone="WHE2495_E11_f21"
                /tissue_type="Early reproductive apex"
                /dev_stage="Seven week-old plants"
                /lab_host="E. coli XL0LR"
                /clone_lib="Triticum monococcum early reproductive apex
                    cDNA library"
                /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
                    Site_1: EcoRI; Site_2: XhoI; The tissue, total RNA, and
                    poly(A) RNA were prepared from apex at double-ridge stage
                    to terminal-spikelet stage during transition from
                    vegetative state to flower state, a cDNA library was made,
                    and the cDNA clones were in vivo excised at the
                    University of California, Davis (V. Echenique, B. Stamova
                    , J. Dubcovsky). Plasmid DNA preparations and DNA
                    sequencing were performed in the OD Anderson lab (all
                    other authors)."

```

```

BASE COUNT      93 a      202 c      139 g      75 t
ORIGIN
Query Match:          5.0%; Score 75; DB 10; Length 509;
Best Local Similarity 50.3%; Pred. NO. 0.0041;
Matches 239; Conservative 0; Mismatches 230; Indels 6; Gaps 2;

QY 241 ACCGTGGCCACCCCTGTACTCGGTGCACGCCCGGCATCGAGTCCGCGGACACCAAGGAGGCC 300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1 ACCGGCTACTCCCTGCTCGCCACCGCTCGCCATCCCGGACGACCGGCACCATCTTGCC 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 301 CTGGACACAGATCGAGGAGGAGCAGAACAACTCCCAGCAGNAGACCCAGCAGGCCAAGGAG 360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 61 ATGGACATCAACCGCGAGRACIACGAGCTGGGGCTGCCGTGCATCGAGAGGCCGCGCGTG 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 361 GCCAGCGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACTGCAGGGCCAGATGGTG 420
    || || || || || || || || || || || || || || || || || || || || ||
DB 121 CGGCACAAGATCGGACTTCGGGAGGGCGCGCGCTGCGGTCTCGACGGCGCTGCTGGAG 180
    || || || || || || || || || || || || || || || || || || || || ||

QY 421 CACGAGGCCATCAGCCCCCGCACCCCTGAAGCCCTGGTGAAGGTGATCGAGGAGAGAGGCC 480
    || || || || || || || || || || || || || || || || || || || || ||
DB 181 GACGAGGCCAACCCAGGCACCTTCGACTTCGTGCGACGCCGACAGGACCAACTAC 240
    || || || || || || || || || || || || || || || || || || || || ||

QY 481 TTCAGCCCCGAGGIGATCCCATGTTACCGCCCTTGAGCGAGGGCGCCACCCCGGAGAC 540
    || || || || || || || || || || || || || || || || || || || || ||
DB 241 CTCAACTACGACGACCGCTCATGAGCTCGTCAAGGTGCGGGGGCCCTCCTCGGCTACGAC 300
    || || || || || || || || || || || || || || || || || || || || ||

QY 541 CTGAACACGATGTGAACACCGTGGGGGCCACCGCCCGCCATGCAGATGCTGAAGGAC 600
    || || || || || || || || || || || || || || || || || || || || ||
DB 301 ---AACACGCTCTGGAACGGCTCCGTGCTGCTCCCGCCGACGCCCCCATGCGCAAGTAC 357
    || || || || || || || || || || || || || || || || || || || || ||

QY 601 ACCATCAACGAGGAGGCCGGCGAGTGGGACCGCCTGCACCCCGTGCAGGCCGGC---CCC 657
    || || || || || || || || || || || || || || || || || || || || ||
DB 358 ATCCGCTACTACCGCGACITCGTGCTCGAGCTCAACAAGGCCCTCGCGCGCGACCGAGCG 417
    || || || || || || || || || || || || || || || || || || || || ||

QY 658 GTGGCCCCCGGCAGATGCGGACACCCCGCGGCGAGGACATCGCGCGGCGCCACCA 712
    || || || || || || || || || || || || || || || || || || || || ||
DB 418 GTCGAGATCTGCCAGCTCCCGCTCGGCGAGCGCATCACCCCTCTGCCCGCGGCGCCA 472
    || || || || || || || || || || || || || || || || || || || || ||

```

RESULT 10
 CD496623/C
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 FEATURES
 source

/strain="Salinas river, CA"
 /db_xref="taxon:69293"
 /clone="CDA23-C11"
 /sex="mixed male and female"
 /tissue_type="heads and internal organs combined"
 /dev_stage="adult"
 /clone_lib="SHGC-CDA"

/note="Vector: lambda ZAP Express/pBK-CMV; Site:1: EcoRI (5' adaptor); Site:2: XhoI (3' linker primer); The mixed organ cDNA library was generated using the ZAP-cDNA method by Stratagene. First strand cDNA synthesis was primed with a 50 bp linker primer containing an oligo dt sequence preceded by a synthetic XhoI site. 5 prime adaptors were used containing an EcoRI cohesive end. The finished cDNAs were inserted in to the ZAP express vector unidirectionally in the sense orientation with respect to the lacZ promoter of pBK-CMV. An amplified library was prepared from approximately 3 million primary clones in the lambda ZAP Express vector. In vivo excision was then used to generate individual pBK-CMV plasmid clones for EST sequencing."

BASE COUNT 186 a 422 c 366 g 196 t

ORIGIN

Query Match 4.9%; Score 74.4; DB 14; Length 1170;
 Best Local Similarity 48.1%; Pred. No. 0.0057;
 Matches 275; Conservative 0; Mismatches 29; Indels 6; Gaps 2;

QY 856 CCGAAGGAGCGCTTCGGGAGCTACGTGGAGCGGCTTCTCAAGACCGCTGC---GGGCGGAG 912
 Db 650 CTCACGGCTCGCTGGGGGACTACGGCGCGGGCTACGGGAGCGCGCTGCTCAACGGGCTCG 591
 QY 913 CAGGCCACACGAGGACGCTGAAGAAGTGGATGACGAGACCGCTGCTGCTGCGGAGAC 972
 Db 590 CTGGGGGACTACGGCGCGGGGACGCGCTGCTCAACGGCTCGCTGGGGGACTACGGGCTAC 531
 QY 973 CCGGACATGACACCATCTCGCGGCTCTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1032
 Db 530 GCGGACGCGCTGCTCAAC---GGCTCGTGGGGGACTACGGCTACGGCGCGCGCGCGCTC 474
 QY 1033 ACGGCTGCCAGGCGCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1092
 Db 473 AAGGCTCGCTGGGGGACTACGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 414
 QY 1093 AGCCAGGCGCAACAGCGCTGAACATCATGATGACAGAGAGCAACITCAAGGGCGCGCGCG 1152
 Db 413 GCGGCGGACGCGCGCTGCTCAACGGCTGCTGGGGGACTACGGCTACGGCGGACGCGCGCTG 354
 QY 1153 AACGTCAAGTGTCTCAACTGGGCGCAAGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCG 1212
 Db 353 AACGGCTCGTGGGGGACTACGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 294
 QY 1213 CGCAAGAAGGCTGCTGGAGTGGCGGCAAGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCG 1272
 Db 293 GCGGACTACGGCGCGCGCTACGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 234
 QY 1273 CGCCAGGCGCAACTTCTGGGGAAGATCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1332
 Db 233 GCGTACGGCGCGCGCTGCTCAACGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 174
 QY 1333 CTGCAGAACCGGACGAGCG 1392
 Db 173 CTGCTCAACGGCTCGCTGGGGGACTACGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 114
 QY 1393 TTCCGCTTCGAGGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1424
 Db 113 CTGGGCGGACTAAGGTACTCCACGATACCAAA 82

RESULT 11

AY103647

LOCUS

DEFINITION AY103647 2598 bp mRNA linear HTC 16 OCT-2002
 Zea mays P00142084 mRNA sequence.

ACCESSION

AY103647.1 GI:21206725

VERSION

HTC.

KEYWORDS

SOURCE

ORGANISM

Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAC

clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 2598)

Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,

Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.

Maize Mapping Project/DuPont Consensus Sequences for Design of

Overgo Probes

Unpublished (2002)

2 (bases 1 to 2598)

Coe,E.H.

Direct Submission

Submitted (25-APR-2002) Maize Mapping Project, University of

Missouri, Columbia, MO 65211, USA

If you are interested in getting corresponding physical clones,

these are publicly available from ZmDB and may be found by BLAST

searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,

www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the

maize cDNA sequences is either Virginia Walbot, Stanford or Pat

Schnable, Iowa State, then clones may be requested from ZmDB:

www.zmdb.iastate.edu.

Location/Qualifiers

1..2598

/organism="Zea mays"

/mol_type="mRNA"

/db_xref="MaizeDB:638378"

/db_xref="taxon:4577"

/clone_lib="Maize Mapping Project/DuPont Consensus

Library"

/note="this sequence is part of a project of EST

assemblies resulting from the application of public

contigs to seed DuPont contigs; this resource was

assembled by DuPont as part of a collaboration for the

overgo addressing of BACS in conjunction with the Maize

Mapping Project"

BASE COUNT 490 a 907 c 795 g 406 t

ORIGIN

Query Match 4.9%; Score 74.4; DB 11; Length 2598;

Best Local Similarity 43.7%; Pred. No. 0.0062;

Matches 620; Conservative 0; Mismatches 756; Indels 42; Gaps 5;

QY 99 CCTGTGTGGCGCCAGCGCGAGCTGCAGGGCTTCGCCCTGACCGCGCGCTGCTGGAGAC 158

Db 712 CCTCATCAGGGCGCGCGCCCAACCGCGAGCGCGCTGACCGGGAGAGGTGGACGC 771

QY 159 CGCGGAGGGCTGCAAGCAGATCATGAAGCAGCTGCAGCGCGCGCTGCAGACCGGACCGA 218

Db 772 CGCGGAGGGCTTCAAGATCGCGCGCATCGAGGGCGCGCTTCTTCAAGCTCAACCCCAAGGA 831

QY 219 GGAGCTGGCAGCCCTGACACACCGTGGCCACCGCTGACTGCGTGCACCGCGCGCATCGA 278

Db 832 GGGCTCGCCATCGTCAACGGCAC-----GTCCGTGGGTCCCGGCTCGC 876

QY 279 GGTCCGGCACACCAAGAGGGCGCTGGACAAAGATCGAGGAGGAGGAGAACAGTCCACCA 338

Db 877 GGCACCGCTGATGTACGACGCCAACGCTGCGCGCTGCTGCGGAGGCTCTGTCGCGCGCT 936

QY 339 GAAGACCCAGCAGGCCCAAGGAGGCCCGACCGCAAGTGCAGCCAGACTACCCCATCGTGCA 398

Db 937 CTCTCGGAGGTCAAGACGGCAAGCCCGAGTACAGGGACCCACCTGACCCACAGCTGAA 996

QY 399 GAACCTGCGAGGGCCAGATGTTGCACCGAGGCCCATCAGCCCCCGACCCCTGACCGCTGGGT 458

Db 997 GCACCCCGGGGTCCATCGAGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTC 1056

QY 459 GAAGTGTATCGAGGAGAGGCGCTTCAGCCCCCGAGGTCATCCCATGTTTACCGCCCTGAG 518

Db 1057 CTTCAATGAAGCAGGGCCAAAGAAATTGAACGAGCTGGAGCCGCTGCTGAAGCCCAAGCAGGA 1116
QY 519 CGAGGGGCGCCACCCCGCAGGACCTGAACACGATGTGA---ACACCGTGGCGGCGACCA 575
Db 1117 CAGGTACGGCTCCGCAGCTCGCCGCACTGGCTGGCCGCCAGATCGAGGTCATCGGCC 1176
QY 576 GGCCGCGATGCAGATGCTGAAGGACACCAATCAACGAGGAGGCGCGGAGTGGGAGCGCT 615
Db 1177 CGCCACCAAGTCCATCGAGCGCGAATCAACTCCGTAACGACACCGCGTCAATGAGCT 1216
QY 636 GCACCCGCTGCAGGCGCGCCCGCTGGCCCGCGGCGCAATCGCGAGCCCGCGG -GGAG 692
Db 1237 CCACCGCGCAAGCGGCTGCACGCGGCAACTTCCAGGGCAACCCCATCGCGGTGTCAT 1296
QY 693 CGACATCGCGCGCGCCACCGACACCCCTGCAGGAGCAGATCGCCCTGATGACCAAGACCC 752
Db 1297 GGACAAACCGCGCCCTCGGCATCGCCACATCGGCACAGCTCATGTTCGGCASTTCCGA 1356
QY 753 CCGCGTCCCGTGGCGGACATCTACAGCGGTGGATCATCTCTGGGCTGAACAGATCGT 812
Db 1357 GCTGCTACAGGATCTTACAAACACCGGCTCACCTCCACCTGGCGGCGAGCCGCAACCC 1416
QY 813 CGCGATGTACAGCCCGCTGAGCATCTCTGGACATCGGCCAGGCGCCCAAGGAGCCCTCCG 872
Db 1417 CAGCCTGGACTACGGCTCAAGGGCAGCGAGATCGCCATGGCTCTACTGCTCCGAGCT 1476
QY 873 CGACTACGTGGACCGCTTCTTCAAGACCCCTGCGCGCGGAGCAGGCCACCCAGGAGTGAA 932
Db 1477 CCAGTACCTGG---GCACCCCATCATCAACACCGCTGCAGAGCGCGGAGCAGCACAACA 1533
QY 933 GAAGTGGATGACCGAGACCGCTGCTGGTGCAGAACCGCAACCCCGACTGCAAGACCATCT 992
Db 1534 GGACGTGAATCCCTCGGCTCGTCTCGGCCAGGAGACCGCGGAGCGGATCGACATCT 1593
QY 993 GCGCGCTCTCGGCGCGCGCCACCCCTGGAGGAGATGATGACCGGCTGCCAGGCGTGG 1052
Db 1594 GAAGCTCATGTCTGCCACCTACATCTGTGGCGCTGTGCCAGCGCGTGCACCTGGCGACCT 1653
QY 1053 CGGCGCGCGCCACAGCGCGCGCTGCTGGCGGAGCGGATGAGCCAGGCCCAACAGCGTGA 1112
Db 1654 CGAGGAGAACATCAAGCGCTGCTGAAGAACACCGTGACCCAGGTGGCCCAAGAGGTGCT 1713
QY 1113 CATCATGATGCAGAGAGCAACTTCAAGGCGCGCGCCGCAAGGTCAGTGTCTGAAGT 1172
Db 1714 GACCATGAACCCCTCGGCGGAGCTCTCCAGCGCGCGCTTCAGCGAGAGGAGGTGATCAG 1773
QY 1173 CGGCAAGAGGGCGCACATCGG-----CAAGAACTGCGCGCGCGCGG 1214
Db 1774 CGCCATCGAGCGCGGAGCGCGTGTTCAGTACGGGAGGAGCGCGGCGAGCGTGGC 1533
QY 1215 CAAGAAGGGGTGCTGGAAGTTCGGCAAGGAGGGCCACAGATGAAGGACTGCACCGAGCG 1274
Db 1834 GCTGATGCAGAAAGCTGCGCGCGCTGCTGGTGGACACCGCTCAGGACGCGGAGCGGGA 1893
QY 1275 CCAGGCGCAAGTCTCTGGCAAGATCTGGCGCCAGCCCAAGGGCGCGCGCGCAATTCCT 1334
Db 1894 CCGGAGCGCTCCGTCTTCCAAAGATCACAGCTTCGAGGAGGAGTCCCGCGGATCT 1953
QY 1335 GCAGAAACCGGAGCGAGCGCGCGCGCTGCCAGCGTGCACCGCGCGCGCGCGGAGAGCTT 1394
Db 1954 GCGCCAGGAGGTGGAGCGCGCGCGCTGGCGGTGCGGAGGGCAACCGCGCGCGGAGAA 2013
QY 1395 CCGCTTCGAGAGAGACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1454
Db 2014 CCGGATCGCGGAGAGCGCGGTCTTCCCGCTGTACCGCTTCTGCGGAGGAGGCTCGGCG 2073
QY 1455 CGAGCCCGCTGACCGCGCTGCGAGCGCTGTTCGCGAGCG 1492
Db 2074 CGTGTTCGAGCGCGGAGAGGCTCAAGTCCCGCGCGG 2111

RESULT 12
CB870888

LOCUS CB870888 649 bp mRNA linear EST 23-APR-2003
DEFINITION HC15J04w HC Hordeum vulgare subsp. vulgare cDNA clone HC15J04
5-PRIME, mRNA sequence.
ACCESSION CB870888
VERSION CB870888.1 GI:30072868
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. vulgare
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Scrophophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
; Triticeae; Hordeum.
1 (bases 1 to 649)
Varshney,R.K., Zhang,H., Burton,R., Stein,N., Landridge,P. and
Graber,A.
Barley ESTs from coleoptile tissue
Unpublished
Contact: Stein Nils
Molecular Markers Group, Department Genbank
Institute of Plant Genetics and Crop Plant Research (IPK)
Corrensstr. 3, 06466, Gatersleben, Germany
Tel: 039482-5522
Fax: 039482-5595
Email: stein@ipk-gatersleben.de
Insert Length: 649 Std Error: 0.00
Plate: 15 row: 3 column: 4
Seq primer: T7.
FEATURES
Location/Qualifiers
1..649
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/strain="cv tokak"
/db_xref="GABI:549289"
/db_xref="taxon:112509"
/clone="HC15J04"
/tissue_type="Root"
/dev_stage="3 week old"
/clone_lib="HC"
/note="6 and 10 hour drought stress by placing plants on
moist paper (75% rel. humidity) in light"
BASE COUNT 128 a 253 c 175 g 92 t
ORIGIN

Query Match 4.9%; Score 74.2; DB 14; Length 649;
Best Local Similarity 47.6%; Pred. No. 0.0058;
Matches 262; Conservative 0; Mismatches 304; Indels 6; Gaps 2;

QY 124 GAGGGCTTCGCCCTGAACCCCGCGCTGCTGGAGACCGCGGAGGCTGCAAGCAGATCAIG 183
Db 45 GAGATCACCGCCAAACCCCATGGAACCTGATGACGACGTGCGCGGAGGCGCCAGTIC 104
QY 184 AAGCAGCTGCAGCCCGCGCTGCAGACCGGCGACCGAGGAGCTCCGAGCTGTACACACC 243
Db 105 CTCAACATGCTGCTCAAGCTCATCGCGCCAAAGAACCATGGAGTGGCGTCTACAGG 164
QY 244 GTGGCCACCGCTGACTGCTGCAGCGCGCGCATCGAGTCCGCGACACAGGAGGCGCTG 303
Db 165 GGCTACTCCCTGCTCGCCACCGCGCTCGCCATCCCGACGACCGCACCATCTTGGCCATG 224
QY 304 GACAAGATCAGGAGGAGGAGGAGAACAAAGTCCACAGCAGAGACCCAGGAGGAGGCGC 363
Db 225 GACATCAACCGCGAGAACTACGAGCTGGGCTGCCGTGCAATCGAAGGCGCGGTGGCG 284
QY 364 GACGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGTGCAC 423
Db 285 CACAAGATCGACTTCGCGAGGCGCGCGCTCCCGCTCCCTGGACGCGCTCTCGAGGAC 344
QY 424 CAGGCCATCAGCCCGCGCGCGCTGACCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 483
Db 345 GAGGCCAACCAGCGACCTTCGACTTCGCTTCGTTGGAGGCGCGACAGGAGGAGGAGGAG 404
QY 484 AGCCCGGAGGTGATCCCGATGTTTACCGCGCGCTGAGCGAGGCGCGCGCGCGCGCGCG 543
Db 405 AACTACCAGGCGCGCTCATGAAGCTCGTGAAGGTCGCGGCGCGCTCCCTCGGCTACGAG --- 461

West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: andersaspw.usda.gov

Sequences have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: Stratagene SK primer.

FEATURES

source
1..562
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WHE2321_F-2_K23"
/tissue_type="Spike before anthesis"
/dev_stage="Adult plant"
/lab_host="E. coli SOLR"
/clone_lib="Wheat pre-anthesis spike cDNA library"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site_1: EcoRI; Site_2: XhoI; Plants were grown in the
greenhouse. Whole spike with awns trimmed. White, green
and yellow anther were collected and total RNA, and
poly(A) RNA were prepared, a cDNA library was made, and
the cDNA clones were in vivo excised to give pBluescript
phagemids in the TJ Close lab (Choi, Close, Fenton) at
the University of California, Riverside. Plasmid DNA
preparations and DNA sequencing were performed in the CD
Anderson lab (all other authors)."

BASE COUNT 108 a 215 c 154 g 85 t

Query Match 4.9%; Score 73.8; DB 10; Length 562;
Best Local Similarity 49.0%; Pred. No. 0.0066;
Matches 257; Conservative 0; Mismatches 262; Indels 6; Gaps 2;

QY 191 TGCAGCCCGCCCTGCAGACCGCGGACCGGAGGAGCTGCGGAGCCCTGTACAAACACCGTGGCCA 250
DB 35 TGCTGCTCAAGCTCATCGGCGCCCAAGAGACCATGGAGATCGGCGTCTACACCGGCTACT 94
QY 251 CCTGTACTGCGTGACGCGCGGCTCGAGGTTCCGCGACACCAAGGAGGCGCCCTGGACAAGA 310
DB 95 CCTGTCTGCCACCGCGCTCGCCATCCCGGACGAGGCCACCATCTTGCCCATGACATCA 154
QY 311 TCGAGGAGGAGCAGACAAAGTCCCAAGAGAGACCAAGGAGGAGGAGGAGGAGGAGGAGGAG 370
DB 155 ACCGCGAGAACTACTATCTGCGGCTGCGTGCATCGAGAGGCGCGGCTGGCGCAAGA 214
QY 371 AGGTAGGCGAGAACTACCCCATCTGTCAGAACTTCAGGCGCCAGATGTTGCACAGGCGCA 430
DB 215 TCGACTTCGCGGAGGCGCGCGGCTGCGGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 274
QY 431 TCAGCCCGCCGACCCCTGAACCGCTGGGTGAAGGTGATCGAGGAGGAGGAGGAGGAGGAGGAG 490
DB 275 ACCAGGAGACCTTCGACTTCGTTCTGCGAGCGCGCAAGGAGAACTACCTCAACTACG 334
QY 491 AGGTATCCCCCAATGTTACCGCCCTGAGCGAGGCGCGCCACCCCGGAGGAGGAGGAGGAGGAG 550
DB 335 ACGAGCGGCTCATGAGCTCGTCAAGGTCGCGGCGCTCTCGGCTACGAC---AACACCG 391
QY 551 TGTGAACACCGTGGCGCGCCACAGCGCCGCTGAGATGCTGAAGGAGACACCATCAAGG 610
DB 392 TCTGGAACGGCTCCGTGCTGCTCCCGCGAGCGCCCGCATGCGGAAGTACATCGGCTACT 451
QY 611 AGGAGGCGCGGAGTGGGACCGCCTGCACCCCGTGCAGGCGCGC---CCGTCGCGCGG 667
DB 452 ACCGCGACTTCGTTCTCGAGCTCAACAAAGGCGCTCGCGCGGAGGAGGAGGAGGAGGAGGAG 511
QY 668 GCCAGATGCGCGACCCCGCGGAGGAGCATCGCGCGCGGCGGCGCA 712
DB 512 GCCAGTCCCGTCTGGGACGGCATCACCCCTCTGCGCGCGGCGCA 556

RESULT 15

BJ261281 694 bp mRNA linear EST 08-APR-2002
BJ261281 Y. Ogihara unpublished cDNA library, Wh_h Triticum
aestivum cDNA clone whh2c03 3', mRNA sequence.

BJ261281
BJ261281.1 GI:20082874

EST.
Triticum aestivum (bread wheat)
Triticum aestivum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.

i (bases 1 to 694)
Ogihara, Y. and Murai, K.

Expressed genes in Triticum aestivum
Unpublished

Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

Location/Qualifiers
1..694

/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="whh2c03"

/tissue_type="Spike at heading date"
/dev_stage="Feekes' scale 10.5"
/clone_lib="Y. Ogihara unpublished cDNA library, Wh_h"
BASE COUNT 128 a 164 c 256 g 145 t 1 others

Query Match 4.9%; Score 73.2; DB 12; Length 694;
Best Local Similarity 50.9%; Pred. No. 0.0085;
Matches 227; Conservative 0; Mismatches 213; Indels 6; Gaps 2;

QY 270 CGGATCGAGGTCGCGGACACCAAGGAGGCGCTGGACAAGATCGAGGAGGAGGAGGAGGAGGAG 329
DB 600 CGCCATCCCGACGACGACCATCTTTGGCCATGGACATCAACCGCGAGAACTACGAGCT 541
QY 330 GTCCAGCAGAGACCCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 369
DB 540 GGGGCTGCGTGCATCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 481
QY 390 CATCTGCAGAACCTGCAGGCGCGGAGATGTTGCACCAAGGAGGAGGAGGAGGAGGAGGAGGAG 449
DB 480 GCGGCTGCGGCTGCTGGACGCGCTGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 421
QY 450 CGCCTGGTGAAGCTGATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 509
DB 420 CGTCTGCTGGAGCGCGGACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 361
QY 510 CGCCCTGAG 569
DB 350 CGTCAAGCTCGGCGGCGCTCCTCGGCTACGAC---AACACGCTCTGGAACGCGCTCCGCTCGT 304
QY 570 CCACAGGCGCGCCATCGAGATGCTGAAGGAGACACCAATCAACGAGGAGGAGGAGGAGGAGGAGGAG 629
DB 303 GCTCCCGCGCGAGCGCGCGGATGCGCAAGTACATCGGCTACTACCGCGGAGCTTCGTCCTCGA 244
QY 630 CGGCTGACCCCGTGCAGGCGCGGCG---CCCGTGGCGCGCGCGGCGGAGATGCGGAGGAGGAGGAGGAG 686
DB 243 CCTCAACAGGCGCTCGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 184
QY 687 CGGAG 712
DB 183 CGGATCACCCCTCTGCGCGCGCGGCGCA 158

Search completed: September 22, 2003, 22:50:20
Job time : 3445.15 secs

PT Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV
PT Env polypeptide and the polypeptide useful for immunizing a mammal
PT especially human against HIV
XX
PS Claim 2; Page 93; 113pp; English.
XX
CC Expression cassettes comprising a polynucleotide encoding antigenic
CC type C human immunodeficiency virus (HIV) Gag or Env polypeptides are
CC useful in DNA immunization, generation of packaging cell lines and
CC production of Gag- and/or Env-containing proteins. Synthetic Env and Gag
CC expression cassettes exhibit increased potency for induction of
CC cytotoxic T-lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1
CC self-assemble into non-infectious virus-like particles which are used as
CC a matrix for the proper presentation of an antigen entrapped or
CC associated to the immune system of the host.
XX
SQ Sequence 1509 BP; 320 A; 556 C; 472 G; 161 T; 0 other;

Query Match 100.0%; Score 1509; DB 21; Length 1509;
Best Local Similarity 100.0%; Pred. No. 2.4e-179;
Matches 1509; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGCGCGCGCCAGGAGCACTACATGCTGAAGCACTGGTGGTGGCCAGCCGGAG 60
DB 1 ATGGGCGCGCGCCAGGAGCACTACATGCTGAAGCACTGGTGGTGGCCAGCCGGAG 60
QY 61 CTGCGCCCGCGCGCCAGGAGCACTACATGCTGAAGCACTGGTGGTGGCCAGCCGGAG 120
DB 61 CTGCGCCCGCGCGCCAGGAGCACTACATGCTGAAGCACTGGTGGTGGCCAGCCGGAG 120
QY 121 CTGGAGGGCTTGGCCCTGAACCCCGGCTGCTGGAGACCGCCGAGGCTGCAAGCAATC 180
DB 121 CTGGAGGGCTTGGCCCTGAACCCCGGCTGCTGGAGACCGCCGAGGCTGCAAGCAATC 180
QY 181 ATGAAGCAGCTGCAGCCCGCCCTGCAGACCGGCGCAGGAGCTGCGCAGCCCTGTACAAC 240
DB 181 ATGAAGCAGCTGCAGCCCGCCCTGCAGACCGGCGCAGGAGCTGCGCAGCCCTGTACAAC 240
QY 241 ACCGTGGCCACCCCTGTACTCGCTGCACCGCGGCTCGAGGTCCGGACACCAAGAGGCC 300
DB 241 ACCGTGGCCACCCCTGTACTCGCTGCACCGCGGCTCGAGGTCCGGACACCAAGAGGCC 300
QY 301 CTGGACAAGATCGAGGAGGAGCAGAACAAAGTCCAGAGAGAGCCAGCAGCGCAAGGAG 360
DB 301 CTGGACAAGATCGAGGAGGAGCAGAACAAAGTCCAGAGAGAGCCAGCAGCGCAAGGAG 360
QY 361 GCGGACGGCAAGGTGAGCCAGAACTACCCCATCTGTGAGAACTGAGGCGGAGATGGTG 420
DB 361 GCGGACGGCAAGGTGAGCCAGAACTACCCCATCTGTGAGAACTGAGGCGGAGATGGTG 420
QY 421 CACGAGGCGATCAGCCCGCCAGCCCTGAAGCCCTGGTGAAGGTGATCCAGGAGAGGCC 480
DB 421 CACGAGGCGATCAGCCCGCCAGCCCTGAAGCCCTGGTGAAGGTGATCCAGGAGAGGCC 480
QY 481 TTCAGCCCGAGGTGATCCCATGTTACCCGCGCTGAGGAGGCGCCACCCCGCAGGAG 540
DB 481 TTCAGCCCGAGGTGATCCCATGTTACCCGCGCTGAGGAGGCGCCACCCCGCAGGAG 540
QY 541 CTGAACACGATGTTGAACACCGCTGGGCGCCACCAAGCCGCTGAGATGCTGAAGGAG 600
DB 541 CTGAACACGATGTTGAACACCGCTGGGCGCCACCAAGCCGCTGAGATGCTGAAGGAG 600
QY 601 ACCATCAACGAGGCGCGGAGTGGGAGGCGCTGACCCCGCTGAGGCGCGCGCCCGG 660
DB 601 ACCATCAACGAGGCGCGGAGTGGGAGGCGCTGACCCCGCTGAGGCGCGCGCCCGG 660
QY 661 GCGCCCGCGGAGATGCGGAGACCCCGCGGAGGAGATGCGGCGCGCCACCAAGAGG 720
DB 661 GCGCCCGCGGAGATGCGGAGACCCCGCGGAGGAGATGCGGCGCGCCACCAAGAGG 720
QY 721 CAGGAGCAGATGCGCTGATGACCAAGACCCCGCCCTGCGGCTGCGGAGATCTACAAG 780
DB 721 CAGGAGCAGATGCGCTGATGACCAAGACCCCGCCCTGCGGCTGCGGAGATCTACAAG 780

QY 781 CGGTGGATCATCTGGGCTGAAACAAGATCTGCGGACTAGCTGGACCGCTCTTCAAGACC 840
DB 781 CGGTGGATCATCTGGGCTGAAACAAGATCTGCGGACTAGCTGGACCGCTCTTCAAGACC 840
QY 841 GACATCCGCGCGCGCGCGCGCGCGCTTCCGCGACTAGCTGGACCGCTCTTCAAGACC 900
DB 841 GACATCCGCGCGCGCGCGCGCGCGCTTCCGCGACTAGCTGGACCGCTCTTCAAGACC 900
QY 901 CTGCGCGCGCGCGCGCGCGCGCGCTGAGAACTGCTGAGACCGGAGACCGCTCTGCTGG 960
DB 901 CTGCGCGCGCGCGCGCGCGCGCGCTGAGAACTGCTGAGACCGGAGACCGCTCTGCTGG 960
QY 961 CAGAACGCGCAACCGCGACTGCAAGACCATCTGCGCGCTCTGCGCGCGCGCGCGCGCTG 1020
DB 961 CAGAACGCGCAACCGCGACTGCAAGACCATCTGCGCGCTCTGCGCGCGCGCGCGCGCTG 1020
QY 1021 GAGGAGATGATGACCGCGCTGCCAGGCGGTGGGCGCGCGCGCGCGCGCGCGCGCTG 1080
DB 1021 GAGGAGATGATGACCGCGCTGCCAGGCGGTGGGCGCGCGCGCGCGCGCGCGCGCTG 1080
QY 1081 GCCAGGCGATGAGCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 1140
DB 1081 GCCAGGCGATGAGCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 1140
QY 1141 GGCG 1200
DB 1141 GGCG 1200
QY 1201 TGCG 1260
DB 1201 TGCG 1260
QY 1261 GACTGCACCGAGCG 1320
DB 1261 GACTGCACCGAGCG 1320
QY 1321 CCGGCAACTTCTGCGAGAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
DB 1321 CCGGCAACTTCTGCGAGAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
QY 1381 CCGGCGAGAGCTTCCGCTTCGAGGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1440
DB 1381 CCGGCGAGAGCTTCCGCTTCGAGGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1440
QY 1441 CGCGAGCGCTACCG 1500
DB 1441 CGCGAGCGCTACCG 1500
QY 1501 AGCCAGTAA 1509
DB 1501 AGCCAGTAA 1509

RESULT 2
AAL44549
ID AAL44549 standard; DNA; 1509 BP.
XX
AC AAL44549;
XX
DT 08-NOV-2002 (first entry)
XX
DE HIV-1 p55gag polypeptide coding sequence 2.
XX
KW HIV; ds; vaccine; gene; immune response; microparticle;
KW adsorbent surface; poly(alpha-hydroxy acid); polyhydroxy butyric acid;
KW polycaprolactone; polyorthoester; polycyanoacrylate; detergent;
KW submicron emulsion; viral infection; bacterial infection;
KW parasitic infection; HIV-1 p55gag polypeptide.
CS Human immunodeficiency virus type 1.
XX
PN WO200226209-A2.

XX 04-APR-2002.
PD
XX 28-SEP-2001; 2001WO-US30540.
PF
XX 28-SEP-2000; 2000US-236105P.
PR
XX 30-AUG-2001; 2001US-315905P.
PR
XX (CHIR) CHIRON CORP.
PA
XX O'hagan D, Otten G, Donnelly JJ, Polo JM, Barnett S, Singh M;
PI Ulmer J, Dubensky TW;
PI WPI; 2002-519084/55.
XX
XX A microparticle to which a biologically active macromolecule is
PT adsorbed, for use as a vaccine composition to treat viral, bacterial or
PT parasitic infections, comprises a polymer microparticle, a detergent
PT and a submicron emulsion -
PT
XX Claim 72; Fig 2; 100pp; English.
PS
XX The invention relates to a method of raising an immune response in a host
CC animal. The method of the invention comprises administering a
CC microparticle that has an adsorbent surface to which a first biologically
CC active macromolecule (e.g. a nucleic acid) has been adsorbed. The
CC microparticle comprises a polymer microparticle of poly(alpha-hydroxy
CC acid), a polyhydroxy butyric acid, a polycaprolactone, a polycrthoester,
CC a polycyanoacrylate, a detergent, and submicron emulsion. The method/
CC microparticle of the invention is useful for immunising a host animal
CC against viral, bacterial or parasitic infections. The present DNA
CC sequence encodes a HIV-1 p55gag polypeptide.
XX
SQ Sequence 1509 BP: 320 A: 556 C: 472 G: 161 T: 0 other;

Query Match 100.0%; Score 1509; DR 24; Length 1509;
Best Local Similarity 100.0%; Pred. No. 2.4e-179;
Matches 1509; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGCCCCCGCCAGCATCTCGCGGGGAGAGCTGGACAAGTGGGAGAGATCCGC 60
Db |||||
QY 61 CTGCGCCCCCGGGGCAAGAGCACTACATGCTGAAGCAACCTGGTGGGCAAGGAG 120
Db |||||
QY 61 CTGCGCCCCCGGGGCAAGAGCACTACATGCTGAAGCAACCTGGTGGGCAAGGAG 120
- Db |||||
QY 121 CTGAGGGCTTGGCCCTGAACCCCGCCCTGCTGGAGACCCGCGAGGGGTGAAGCAGATC 180
- Db |||||
QY 181 ATGAGCAGCTGCAGCCCGCCCTGCAGACCCGAGAGCTGCGAGCTGTACAAAC 240
Db |||||
QY 241 ACCGTGGCCACCCCTGTACTGCTGCACCCCGCATCGAGGTCCGCGACACCAAGGAGGCC 300
Db |||||
QY 301 CTGGACAGATCGAGGAGGAGCAGAACAAAGTCCAGCAGACAGACCCAGAGGCCAAGGAG 360
Db |||||
QY 361 GCCGACGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACTGCAGGGCCAGATGGTG 420
Db |||||
QY 421 CACGAGCCATCAGCCCCCGCCACCTGTACCCGCGCTGAGCGAGGGGCCACCCCGCAGGAC 480
Db |||||
QY 481 TTCAGCCCGGAGGTGATCCCATGTTCACCGCGCTGAGCGAGGGGCCACCCCGCAGGAC 540
|||

Lb 481 TTCAGCCCGGAGGTGATCCCATGTTCACCGCGCTTGAGCGAGGGGCCACCCCGCAGGAC 540
QY 541 CTGAACACAGTGTGAACACCGTGGCGGCCACCAAGTCCGCCATGACATGCTGAAGGAC 600
Db |||||
QY 541 CTGAACACAGTGTGAACACCGTGGCGGCCACCAAGTCCGCCATGACATGCTGAAGGAC 600
Db |||||
QY 601 ACCATCAACGAGGAGGCCGCCGAGTGGGACCGCTGCACCCCGTGCAGGCCGCCCGCTG 660
Db |||||
QY 601 ACCATCAACGAGGAGGCCGCCGAGTGGGACCGCTGCACCCCGTGCAGGCCGCCCGCTG 660
Db |||||
QY 661 GCGCCCGCCAGATGCGCGACCCCGCGGAGCGACATCGCCGCGGCGGCGGCGGCGGCG 720
Db |||||
QY 721 CAGGAGCAGATGCGCTGGATGACGACCAACCCCGCTGCGCGTGGGCGACATCTACAAG 780
Db |||||
QY 721 CAGGAGCAGATGCGCTGGATGACGACCAACCCCGCTGCGCGTGGGCGACATCTACAAG 780
Db |||||
QY 781 CGGTGGATCATCTCGTGGCTGAACAGATCGTGGGATGTACAGCCCGCTGAGCATCTCTG 840
Db |||||
QY 781 CGGTGGATCATCTCGTGGCTGAACAGATCGTGGGATGTACAGCCCGCTGAGCATCTCTG 840
Db |||||
QY 841 GACATCGCCAGGCGCCCAAGGAGCCCTTCGCGGACTAGTGGACCGCTTCTTCAAGACC 900
Db |||||
QY 841 GACATCGCCAGGCGCCCAAGGAGCCCTTCGCGGACTAGTGGACCGCTTCTTCAAGACC 900
Db |||||
QY 901 CTGCGCGCGGAGCGCCACCCAGGACGCTGAAGAACTGATGACCGAGACCCCTGCTGGTG 960
Db |||||
QY 901 CTGCGCGCGGAGCGCCACCCAGGACGCTGAAGAACTGATGACCGAGACCCCTGCTGGTG 960
Db |||||
QY 961 CAGAACGCCAACCCCGACTGCAAGACCATCTGCGCGCTCTCGGCCGCGGCGGCGGCG 1020
Db |||||
QY 961 CAGAACGCCAACCCCGACTGCAAGACCATCTGCGCGCTCTCGGCCGCGGCGGCGGCG 1020
Db |||||
QY 1021 GAGGAGATGATGACCCGCTGCCAGGCGGTGGGCGGCCGCGGCGGCGGCGGCGGCTG 1080
Db |||||
QY 1021 GAGGAGATGATGACCCGCTGCCAGGCGGTGGGCGGCCGCGGCGGCGGCGGCGGCTG 1080
Db |||||
QY 1081 GCGGAGGCGATGAGCCAGGCGCAACAGCGTGAACATCATGATGACAGAGGCAACTTCAAG 1140
Db |||||
QY 1081 GCGGAGGCGATGAGCCAGGCGCAACAGCGTGAACATCATGATGACAGAGGCAACTTCAAG 1140
Db |||||
QY 1141 GCGCCCGCGGCGCAAGTCAAGTCTTCACTTCCGCGAGGAGGCGGCGGCGGCGGCGG 1200
Db |||||
QY 1141 GCGCCCGCGGCGCAAGTCAAGTCTTCACTTCCGCGAGGAGGCGGCGGCGGCGGCGG 1200
Db |||||
QY 1201 TGCGCGCGCGCGCAAGAGGCGTCTGGAAGTGGCGGAGGAGGCGGCGGCGGCGGCGG 1260
Db |||||
QY 1261 GACTGCACCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1320
Db |||||
QY 1261 GACTGCACCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1320
Db |||||
QY 1321 CCGGCGCAACTTCTGCGAGAACCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1380
Db |||||
QY 1321 CCGGCGCAACTTCTGCGAGAACCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1380
Db |||||
QY 1381 CCGCGCGGAGGCTTCCGCTTCAGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1440
Db |||||
QY 1381 CCGCGCGGAGGCTTCCGCTTCAGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1440
Db |||||
QY 1441 GCGGAGCCCTACGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1500
Db |||||
QY 1441 GCGGAGCCCTACGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1500
Db |||||
QY 1501 AGCCAGTAA 1509
Db |||||
QY 1501 AGCCAGTAA 1509
Db |||||

QY 1381 CCGGCGGAGAGCTTCGGCTTCGAGGAGACACCCCGGCCCCCAGAGAGAGGAC 1440
 DB 1381 CCGGCGGAGAGCTTCGGCTTCGAGGAGACACCCCGGCCCCCAGAGAGAGGAC 1440
 QY 1441 CGCGAGCCCTACCGGAGCCCTGACCGCCCTGCCAGCCTGTTCGGCAGCGGCCCTG 1500
 DB 1441 CGCGAGCCCTACCGGAGCCCTGACCGCCCTGCCAGCCTGTTCGGCAGCGGCCCTG 1500
 QY 1501 AGCCAGTAA 1509
 DB 1501 AGCCAGTAA 1509

RESULT 4
 ID ABL39958 standard; DNA: 1509 BP.
 XX ABL39958;
 AC
 XX
 DT 15-MAY-2002 (first entry)
 XX
 DE Synthetic Gag polynucleotide sequence SEQ ID NO:21.
 XX
 KW Human immunodeficiency virus type C; antigenic HIV type C protein;
 KW immunogenic; immunisation: gag; pol; vif; vpr; tat; rev; vpu; env; nef;
 KW immunostimulant; gene therapy; gene; ds.
 XX
 OS Human immunodeficiency virus type C.
 OS Synthetic.
 XX
 PN WO200204493-A2.
 XX
 PD 17-JAN-2002.
 XX
 PF 05-JUL-2001; 2001WO-US21241.
 XX
 PR 05-JUL-2000; 2000US-0610313.
 XX
 PA (CHIR) CHIRON CORP.
 PA (UYST-) UNIV STELLENBOSCH.
 XX
 PI Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;
 XX
 DR WPI; 2002-154920/20.
 XX
 CC New polynucleotides encoding antigenic HIV Type C polypeptides, useful
 PT in applications including DNA immunization or generation of packaging
 PT cell lines, particularly in gene therapy -
 XX
 PS Example 1; Fig 6: 233pp; English.
 XX
 CC The present invention describes expression cassettes comprising a
 CC polynucleotide sequence encoding a polypeptide comprising immunogenic
 CC HIV type C polypeptides. The expression cassettes comprise any of the
 CC HIV type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or
 CC Nef (1). (I) have immunostimulant activity and can be used in gene
 CC therapy. The HIV type C polynucleotides are useful in applications
 CC including DNA immunisation, generation of packaging cell lines, and
 CC production of HIV Type C proteins. The polynucleotides are particularly
 CC useful in gene therapy and DNA immunisation applications. ABL39942 to
 CC ABL40054 and ABL06204 to ABL06215 represent sequences used in the
 CC exemplification of the present invention.
 XX
 SQ Sequence 1509 BP; 321 A; 560 C; 470 G; 158 T; 0 other;

Query Match 98.8%; Score 1491.4; DB 24; Length 1509;
 Best Local Similarity 99.3%; Pred. No. 3.7e-177;
 Matches 1498; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 ATGGGCGCGCGCGGAGCATCTCGCGGCGGAGAGCTGGACAGTGGGAGAGATCCGC 60
 DB 1 ATGGGCGCGCGCGGAGCATCTCGCGGCGGAGAGCTGGACAGTGGGAGAGATCCGC 60

QY 51 CTGGCGCCCGGGGCAAGAGCACTACATGCTGAAGCACCTGGTGTGGGCCAGCCGCGAG 120
 DB 61 CTGGCGCCCGGGGCAAGAGCACTACATGCTGAAGCACCTGGTGTGGGCCAGCCGCGAG 120
 QY 121 CTGAGGGGCTTCGCCCTTGAACCGCGGCTGCTGTGGAGACCGCGGAGGGCTGCAAGCAGATC 180
 DB 121 CTGAGGGGCTTCGCCCTTGAACCGCGGCTGCTGTGGAGACCGCGGAGGGCTGCAAGCAGATC 180
 QY 181 ATGAGCAGCTGCAGCCCGCTTCAGACCCGCGACCGGAGAGCTGCGCAGCCTGTACAC 240
 DB 181 ATGAGCAGCTGCAGCCCGCTTCAGACCCGCGACCGGAGAGCTGCGCAGCCTGTACAC 240
 QY 241 ACCGTGGCCACCCCTGACTGCTGTCACGCGCGGCTCGAGGTCGCGCACACCAAGGAGGC 300
 DB 241 ACCGTGGCCACCCCTGACTGCTGTCACGCGCGGCTCGAGGTCGCGCACACCAAGGAGGC 300
 QY 301 CTGGACAAGATCGAGGAGGAGGAGCAACAAGTCCCGAGCAGACACCCAGCAGCCCAAGGAG 360
 DB 301 CTGGACAAGATCGAGGAGGAGGAGCAACAAGTCCCGAGCAGACACCCAGCAGCCCAAGGAG 360
 QY 361 GCCGACGGCAAGGTGAGCCAGAACTACGCCCATCGGTCAGAACCTGCAGGGCCAGTGGTG 420
 DB 361 GCCGACGGCAAGGTGAGCCAGAACTACGCCCATCGGTCAGAACCTGCAGGGCCAGTGGTG 420
 QY 421 CACGAGGCCATCAGCCCCCGCACCTGAACGGCTGGGTGAAGGTGATCGAGGAGAGGCC 480
 DB 421 CACGAGGCCATCAGCCCCCGCACCTGAACGGCTGGGTGAAGGTGATCGAGGAGAGGCC 480
 QY 481 TTCAGCCCGGAGGTGATCCCATGTTTCAACGGCTGAGCGAGGGCGCCACCCCGCAGGAC 540
 DB 481 TTCAGCCCGGAGGTGATCCCATGTTTCAACGGCTGAGCGAGGGCGCCACCCCGCAGGAC 540
 QY 541 CTGAACACGATGTTGAACACCGCTGGGCGGCGACCGCTGCACCGCTGCAGGCGGCCGCTG 600
 DB 541 CTGAACACGATGTTGAACACCGCTGGGCGGCGACCGCTGCACCGCTGCAGGCGGCCGCTG 600
 QY 601 ACCATCAACGAGGAGGCGCGCGAGTGGGACCGCTGCACCGCTGCAGGCGGCCGCTG 660
 DB 601 ACCATCAACGAGGAGGCGCGCGAGTGGGACCGCTGCACCGCTGCAGGCGGCCGCTG 660
 QY 661 GCCCGCGGCGAGTGGCGGACCGCGCGGCGGAGGAGATCGCGGCGGCCACGACATCGCGGCGGCCACGACCCCTG 720
 DB 661 GCCCGCGGCGAGTGGCGGACCGCGCGGCGGAGGAGATCGCGGCGGCCACGACATCGCGGCGGCCACGACCCCTG 720
 QY 721 CAGGAGCAGATCGCTGGATGATGACCAAGATCGTGGGATGTACAGCCCGCTGAGCATCCTG 840
 DB 721 CAGGAGCAGATCGCTGGATGATGACCAAGATCGTGGGATGTACAGCCCGCTGAGCATCCTG 840
 QY 841 GACATCCCGCAGGCGGCCCAAGAGGCCCTTCGCGGACTACGTGGACCGCTTCTTCAAGACC 900
 DB 841 GACATCCCGCAGGCGGCCCAAGAGGCCCTTCGCGGACTACGTGGACCGCTTCTTCAAGACC 900
 QY 901 CTGGCGCGCGAGGAGGCCACCCAGGACGCTGAAGAACCTGGATGACCGGAGACCGCTGCTGGTG 960
 DB 901 CTGGCGCGCGAGGAGGCCACCCAGGACGCTGAAGAACCTGGATGACCGGAGACCGCTGCTGGTG 960
 QY 961 CAGAACGCCAACCCCGACTGCAAGACCATCTCTGCGGCTCTCGGCCCGCGGCCACCCCTG 1020
 DB 961 CAGAACGCCAACCCCGACTGCAAGACCATCTCTGCGGCTCTCGGCCCGCGGCCACCCCTG 1020
 QY 1021 GAGGAGATGATGACCGCTGCCAGGCGCTGGGCGGCCCGGCCACCAAGGCCCGCTGCTG 1080
 DB 1021 GAGGAGATGATGACCGCTGCCAGGCGCTGGGCGGCCCGGCCACCAAGGCCCGCTGCTG 1080
 QY 1081 GCCGAGGCGATGAGCCAGGCCCAAGCGGTGAACATCATGATGACAGAGAGCAACTTCAAG 1140
 DB 1081 GCCGAGGCGATGAGCCAGGCCCAAGCGGTGAACATCATGATGACAGAGAGCAACTTCAAG 1140
 QY 1141 GGCCCCCGGCGGACAGTCAAGTCTTCAACTCGCGGCAAGGAGGCCACATCGCCAGAAC 1200

Db 1141 GGGCCCGCCGCAACGTGAAGTGTCTCAACTCGCGCAAGGAGGGCCACATCGCCCAAGAAC 1200
QY 1201 TGGCGGCGCCCGCCGCAAGGAGGCTGTGGAAGTGGGCAAGGAGGGCCACAGATGAAG 1260
Db 1201 TGGCGGCGCCCGCCGCAAGGAGGCTGTGGAAGTGGGCAAGGAGGGCCACAGATGAAG 1260
QY 1261 GACTGCACGAGAGCGCCAGGCGCAACTTCTGGGCAAGATCTGGCCAGGCGCAAGGAGGCGG 1320
Db 1261 GACTGCACGAGAGCGCCAGGCGCAACTTCTGGGCAAGATCTGGCCAGGCGCAAGGAGGCGG 1320
QY 1321 CCGGCAACTTCTGCAAGAACCGGAGGAGCGCCGCGCCCGCCCGCCCGCCCGCCCGCCCG 1380
Db 1321 CCGGCAACTTCTGCAAGAACCGGAGGAGCGCCGCGCCCGCCCGCCCGCCCGCCCGCCCG 1380
QY 1381 CCGGCGGAGAGCTTCGGCTTCGAGGAGACCAACCCCGCCCGCCCGCCCGCCCGCCCGCC 1440
Db 1381 CCGGCGGAGAGCTTCGGCTTCGAGGAGACCAACCCCGCCCGCCCGCCCGCCCGCCCGCC 1440
QY 1441 CGGAGGCGCTACGGGAGCGCCCTGACCGCCCTGGGCAAGCCTGTTCGGCAGGCGGCGCCCTG 1500
Db 1441 CGGAGGCGCTACGGGAGCGCCCTGACCGCCCTGGGCAAGCCTGTTCGGCAGGCGGCGCCCTG 1500
QY 1501 AGCCAGTAA 1509
Db 1501 AGCCAGTAA 1509

RESULT 5
AAA51626
ID AAA51626 standard; DNA; 1509 BP.
XX
AC AAA51626;
XX
DT 31-OCT-2000 (first entry)
XX
DE HIV codon-optimized synthetic Gag polynucleotide.
XX
KW Gag; expression cassette; antigenic; type C; HIV; Env; synthetic;
KW DNA immunization; packaging cell line; antigen presentation; ss.
XX
OS Human immunodeficiency virus type C strain AF10967.
OS Synthetic.
XX
PN WO200039304-A2.
XX
PD 06-JUL-2000.
XX
PF 30-DEC-1999; 99WO-US31273.
XX
PR 31-DEC-1998; 98US-0114495.
PR 01-SEP-1999; 99US-0152195.
XX
PA (CHIR) CHIRON CORP.
XX
PI Barnett S. Zur Megede J;
XX
DR WPI; 2000-452401/39.
XX
PT Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV
PT Env polypeptide and the polypeptide useful for immunizing a mammal
PT especially human against HIV
XX
PS Disclosure; Page 104; 113pp; English.
XX
CC Expression cassettes comprising a polynucleotide encoding antigenic
CC type C human immunodeficiency virus (HIV) Gag or Env polypeptides are
CC useful in DNA immunization, generation of packaging cell lines and
CC production of Gag- and/or Env-containing proteins. Synthetic Env and Gag
CC expression cassettes exhibit increased potency for induction of
CC cytotoxic T-lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1
CC self-assemble into non-infectious virus-like particles which are used as
CC a matrix for the proper presentation of an antigen entrapped or

CC associated to the immune system of the host.
XX
SQ Sequence 1509 BP; 321 A; 559 C; 471 G; 158 T; 0 other;
Query Match 98.7%; Score 1489.8; DB 21; Length 1509;
Best local Similarity 99.2%; Pred. No. 5.8e-177;
Matches 1497; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
QY 1 ATGGGCGCGCGGCGAGCATCTCCGCGGCGGAGAACGCTGGGAGAGATCCGC 60
Db 1 ATGGGCGCGCGGCGAGCATCTCCGCGGCGGAGAACGCTGGGAGAGATCCGC 60
QY 61 CTGGGCGCGCGGCGAAGAACATACATGTAAGCACCTGGTGTGGCGAGCGCGAG 120
Db 61 CTGGGCGCGCGGCGAAGAACATACATGTAAGCACCTGGTGTGGCGAGCGCGAG 120
QY 121 CTGGAGGGCTTCGGCTGAACCGCGGCTGTGGAGACCGCGAGGGCTGCAAGCAGATC 180
Db 121 CTGGAGGGCTTCGGCTGAACCGCGGCTGTGGAGACCGCGAGGGCTGCAAGCAGATC 180
QY 181 ATGAGCAGCTGCAGCGCGGCTGCAGACCGGCGAGGAGCTGGCAGGCTGTACAC 240
Db 181 ATGAGCAGCTGCAGCGCGGCTGCAGACCGGCGAGGAGCTGGCAGGCTGTACAC 240
QY 241 ACGTGGCGACCCCTGTACTGCTGTCACCGCGGCTGAGGTGCGGACCAAGAGGCC 300
Db 241 ACGTGGCGACCCCTGTACTGCTGTCACCGCGGCTGAGGTGCGGACCAAGAGGCC 300
QY 301 CTGGACAAGATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
Db 301 CTGGACAAGATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
QY 361 GCGGACGCAAGGTGAGCCAGAACTACCCCTGCTGCAGAACCTGAGGGCGGAGATGGTG 420
Db 361 GCGGACGCAAGGTGAGCCAGAACTACCCCTGCTGCAGAACCTGAGGGCGGAGATGGTG 420
QY 421 CACCAGGCCATCAGGCG 480
Db 421 CACCAGGCCATCAGGCG 480
QY 481 TTGAGCGCGGAGGTGATCCCGATGTTACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
Db 481 TTGAGCGCGGAGGTGATCCCGATGTTACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
QY 541 CTGACACAGATGTTGAACACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
Db 541 CTGACACAGATGTTGAACACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
QY 601 ACCATCAACGAGGAGGCG 660
Db 601 ACCATCAACGAGGAGGCG 660
QY 661 GCGCGCGCGCGAGATGCG 720
Db 661 GCGCGCGCGCGAGATGCG 720
QY 721 CAGGAGCAGATGCG 780
Db 721 CAGGAGCAGATGCG 780
QY 781 CGGTGATCATCTGGCGCTGAACAGAGATCGTGGCGCATGTACAGCGCGCGCGCGCGCG 840
Db 781 CGGTGATCATCTGGCGCTGAACAGAGATCGTGGCGCATGTACAGCGCGCGCGCGCGCG 840
QY 841 GACATCG 900
Db 841 GACATCG 900
QY 901 CTGGCGCGCGGAGCG 960
Db 901 CTGGCGCGCGGAGCG 960
QY 961 CAGAACGCGCAACCG 1020


```
Db 961 CAGACGCCAACCCGACIGCAAGACCACTCTGCGGCCCTGGGCCCGCGCCACACCTG 1020
QY 1021 GAGGAGATGATGACCGCTGCCAGGGCGTGGCGGCCCGCCGACACAGCGCGGTGCTG 1080
Db 1021 GAGGAGATGATGACCGCTGCCAGGGCGTGGCGGCCCGCCGACACAGCGCGGTGCTG 1080
QY 1081 GCCGAGCGATGAGCCAGGCCCAACACCGTGACATCATGATGACAGAGAGCACTTCAAG 1140
Db 1081 GCCGAGCGATGAGCCAGGCCCAACACCGTGACATCATGATGACAGAGAGCACTTCAAG 1140
QY 1141 GSCCCCGCGCCCAACGICAACTGCTCAACTGCGGCAAGGAGGCCACATCGCCCAAGAAC 1200
Db 1141 GSCCCCGCGCCCAACGICAACTGCTCAACTGCGGCAAGGAGGCCACATCGCCCAAGAAC 1200
QY 1201 TGCGCGCCCGCCCGCAAGAGGGCTGCTGGAAGTGCGGCAAGGAGGCCACAGATGAAG 1260
Db 1201 TGCGCGCCCGCCCGCAAGAGGGCTGCTGGAAGTGCGGCAAGGAGGCCACAGATGAAG 1260
QY 1261 GACTGACCGGAGCGCCAGGCCCAACTTCCTGGGCAAGATCTGSCCCAGCCACAGGGCGCG 1320
Db 1261 GACTGACCGGAGCGCCAGGCCCAACTTCCTGGGCAAGATCTGSCCCAGCCACAGGGCGCG 1320
QY 1321 CCGCGCAACTTCCTGAGAGACCGGACCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
Db 1321 CCGCGCAACTTCCTGAGAGACCGGACCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
QY 1381 CCGCGGAGAGGCTTCGCGGAGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1440
Db 1381 CCGCGGAGAGGCTTCGCGGAGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1440
QY 1441 CCGGAGCGCTACCGGAGCGCGCTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1500
Db 1441 CCGGAGCGCTACCGGAGCGCGCTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1500
QY 1501 AGCCAGTAA 1509
Db 1501 AGCCAGTAA 1509
```

RESULT 6

AAL44553

ID AAL44553 standard; DNA; 1509 BP.

XX

AC AAL44553;

XX 08-NOV-2002 (first entry)

DE HIV-1 p55gag polypeptide coding sequence 4.

KW HIV; ds: vaccine; gene: immune response; microparticle;
adsorbent surface; poly(alpha-hydroxy acid); polyhydroxy butyric acid;
polycaprolactone; polyorthoester; polycyanoacrylate; detergent;
submicron emulsion; viral infection; bacterial infection;
parasitic infection; HIV-1 p55gag polypeptide.

OS Human immunodeficiency virus type 1.

XX Key

FH Location/Qualifiers

FT variation: replace (282, C)

FT /*tag= a

FT variation: replace (331, T)

FT /*tag= b

FT variation: replace (332, C)

FT /*tag= c

FT variation: replace (549, G)

FT /*tag= d

FT variation: replace (553, T)

FT /*tag= e

FT variation: replace (783, G)

FT /*tag= f

FT variation: replace (816, G)

FT /*tag= g

FT variation: replace (999, T)
FT /*tag= h
FT variation: replace (1002, C)
FT /*tag= i
FT variation: replace (1089, G)
FT /*tag= j
FT variation: replace (1149, G)
FT /*tag= k
FT variation: replace (1158, C)
FT /*tag= l
XX
PN WO200226209-A2.
XX
PD 04-APR-2002.
XX
PF 28-SEP-2001; 2001WO-US30540.
XX
PR 28-SEP-2000; 2000US-236105P.
PR 30-AUG-2001; 2001US-315905P.
XX
PA (CHIR) CHIRON CORP.
XX
PI O'hagan D, Otten G, Donnelly JJ, Polo JM, Barnett S, Singh M;
PI Ulmer J, Dubensky TW;
XX
DR WPI; 2002-519084/55.
XX
PT A microparticle to which a biologically active macromolecule is
adsorbed, for use as a vaccine composition to treat viral, bacterial or
parasitic infections, comprises a polymer microparticle, a detergent
and a submicron emulsion -
XX
PS Claim 72; Fig 5; 100pp; English.
XX
CC The invention relates to a method of raising an immune response in a host
animal. The method of the invention comprises administering a
microparticle that has an adsorbent surface to which a first biologically
active macromolecule (e.g. a nucleic acid) has been adsorbed. The
microparticle comprises a polymer microparticle of poly(alpha-hydroxy
acid), a polyhydroxy butyric acid, a polycaprolactone, a polyorthoester,
a polycyanoacrylate, a detergent, and submicron emulsion. The method/
microparticle of the invention is useful for immunising a host animal
against viral, bacterial or parasitic infections. The present DNA
sequence encodes a HIV-1 p55gag polypeptide.
XX
SQ Sequence 1509 BP; 321 A; 559 C; 471 G; 158 T; 0 other;

Query Match 98.7%; Score 1489.8; DB 24; Length 1509;
Best Local Similarity 99.2%; Pred. No. 5.8e-177;
Matches 1497; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 ATGGGCGCCCGCGCGGCAAGACCTACATGCTGAAGCACCTGGTGTGGCCAGCCGGAG 120
|||||
Db 1 ATGGGCGCCCGCGCGGCAAGACCTACATGCTGAAGCACCTGGTGTGGCCAGATCCGC 60
|||||
QY 51 CTGGCGCCCGCGCGGCAAGACCTACATGCTGAAGCACCTGGTGTGGCCAGCCGGAG 120
|||||
Db 51 CTGGCGCCCGCGCGGCAAGACCTACATGCTGAAGCACCTGGTGTGGCCAGCCGGAG 120
|||||
QY 121 CTGGAGGGCTTCGCCCTGAACCCCGGCCCTGCTGTGAGAGACCCGCGAGGCTGCACAGCATC 180
|||||
Db 121 CTGGAGGGCTTCGCCCTGAACCCCGGCCCTGCTGTGAGAGACCCGCGAGGCTGCACAGCATC 180
|||||
QY 181 ATGAAGCAGCTGCAGCCCGCCCTGCAGACCCGCGAGGAGGAGCTGCCAGCCTGTACAAC 240
|||||
Db 181 ATGAAGCAGCTGCAGCCCGCCCTGCAGACCCGCGAGGAGGAGCTGCCAGCCTGTACAAC 240
|||||
QY 241 ACCGTGGCCACCCCTGTACTGGTGTGACGACCCCGGCATCGAGGTCCGCGACACCCAGGAGGCC 300
|||||
Db 241 ACCGTGGCCACCCCTGTACTGGTGTGACGACCCCGGCATCGAGGTCCGCGACACCCAGGAGGCC 300
|||||
QY 301 CTGGACAAGATCGAGGAGGAGCAGACAAGTCCAGCAGAGACCCAGCAGCCCAAGGAG 360
|||||

Db 301 CTGGACAAGATCGAGGAGGAGCAGAAACAAAGAGCCACGAGAAGACCCAGCAGGGCCCAAGGAG 360

Qy 361 GCCAGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGAGGCCAGATGGTG 420

Db 361 GCGAGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGAGGCCAGATGGTG 420

Qy 421 CACCAGGCCATCAGCCCGCCAGCCCTGAACGCCCTGGTGAAGGTGATCGAGGAGAGGCC 430

Db 421 CACCAGGCCATCAGCCCGCCAGCCCTGAACGCCCTGGTGAAGGTGATCGAGGAGAGGCC 430

Qy 481 TTCAGCCCGGAGGTGATCCCATCTTACCGCCCTGAGCGAGGGCCACCCCGAGGAG 540

Db 481 TTCAGCCCGGAGGTGATCCCATCTTACCGCCCTGAGCGAGGGCCACCCCGAGGAG 540

Qy 541 CTGAACACCATCTCTGAACACCGTGGGGGGCCACCCAGCGCGCCATGCGAGATGCTCAAGGAG 600

Db 541 CTGAACACCATCTCTGAACACCGTGGGGGGCCACCCAGCGCGCCATGCGAGATGCTCAAGGAG 600

Qy 601 ACCATCAACGAGGAGGGCGCGAGTGGGACCGCCCTGCACGCCCTGCAGGCGCGGCGGCTG 660

Db 601 ACCATCAACGAGGAGGGCGCGAGTGGGACCGCCCTGCACGCCCTGCAGGCGCGGCGGCTG 660

Qy 661 GCCCGCGCCAGATGCGGAGACCCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720

Db 661 GCCCGCGCCAGATGCGGAGACCCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720

Qy 721 CAGGAGCAGATCGCCTGGATGACCAAGACCCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780

Db 721 CAGGAGCAGATCGCCTGGATGACCAAGACCCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780

Qy 781 CGGTGGATCACTCTGGGCTGAACAAGATCGTGGGATGTATACAGCCCGGTGAGCACTCTG 840

Db 781 CGGTGGATCACTCTGGGCTGAACAAGATCGTGGGATGTATACAGCCCGGTGAGCACTCTG 840

Qy 841 GACATCCCGCAGGCGCCAAAGGAGCCCTTCGCGACTAGTGAGGACCGCTCTTCAAGACC 900

Db 841 GACATCCCGCAGGCGCCAAAGGAGCCCTTCGCGACTAGTGAGGACCGCTCTTCAAGACC 900

Qy 901 CTGCGCGCGGAGCGCCACCCAGGAGCGCTTCGCGACTAGTGAGGACCGCTCTTCAAGACC 960

Db 901 CTGCGCGCGGAGCGCCACCCAGGAGCGCTTCGCGACTAGTGAGGACCGCTCTTCAAGACC 960

Qy 961 CAGAACGCCAACCCCGACTGCAAGACCATCTCGCGCTCTCGGCCCGCGCCACCGCTG 1020

Db 961 CAGAACGCCAACCCCGACTGCAAGACCATCTCGCGCTCTCGGCCCGCGCCACCGCTG 1020

Qy 1021 GAGGAGATGATGACCGCTGCCAGGGCGTGGCGGCGCCACAGCGCGCGGCTG 1080

Db 1021 GAGGAGATGATGACCGCTGCCAGGGCGTGGCGGCGCCACAGCGCGCGGCTG 1080

Qy 1081 GCCAGGCGATGAGCCAGCCAAACAGCGTGACATCATGATGCAGAGAGCAACTTCAAG 1140

Db 1081 GCCAGGCGATGAGCCAGCCAAACAGCGTGACATCATGATGCAGAGAGCAACTTCAAG 1140

Qy 1141 GGCCCGCGCGCAACGTCAAGTGCTTCAACTGCGGCAAGGAGGCGCCACATCGCCAGAGAC 1200

Db 1141 GGCCCGCGCGCAACGTCAAGTGCTTCAACTGCGGCAAGGAGGCGCCACATCGCCAGAGAC 1200

Qy 1201 TGCCCGCGCGCGCGCAAGGGCTGCTGGAGTGCGGCAAGGAGGAGGAGGAGGAGGAGGAGGAG 1260

Db 1201 TGCCCGCGCGCGCGCAAGGGCTGCTGGAGTGCGGCAAGGAGGAGGAGGAGGAGGAGGAGGAG 1260

Qy 1261 GACTGCACGAGCGCCAGCCAACTTCTGGGCAAGATGTGGCCCGAGCCACAGGCGCG 1320

Db 1261 GACTGCACGAGCGCCAGCCAACTTCTGGGCAAGATGTGGCCCGAGCCACAGGCGCG 1320

Qy 1321 CCGCGCAACTTCTGCAGAACCGCAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380

Db 1321 CCGCGCAACTTCTGCAGAACCGCAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380

Qy 1381 CCGCGCGAGAGCTTCGGCTTCGAGGAGACACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1440

Db 1381 CCGCGCGAGAGCTTCGGCTTCGAGGAGACACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1440

Qy 1441 CCGGAGCCCTACCGCGAGCCCGGAGCCCGGAGCCCGCTGCGGAGCCCTGTCGGCAGGCGCCCTG 1500

Db 1441 CCGGAGCCCTACCGCGAGCCCGGAGCCCGGAGCCCGCTGCGGAGCCCTGTCGGCAGGCGCCCTG 1500

Qy 1501 AGCCAGTAA 1509

Db 1501 AGCCAGTAA 1509

RESULT 7

ABL39972

ID ABL39972 standard; DNA: 1494 BP.

XX ABL39972;

XX 15-MAY-2002 (first entry);

XX Synthetic Gag polynucleotide sequence SEQ ID NO:51.

DE Human immunodeficiency virus type C; antigenic HIV type C protein;

XX immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;

KW immunostimulant; gene therapy; gene; ds.

KW Human immunodeficiency virus type C.

XX Synthetic.

XX WO200204493-A2.

XX 17-JAN-2002.

XX 05-JUL-2001; 2001WO-US21241.

XX 05-JUL-2000; 2000US-0610313.

XX (CHIR) CHIRON CORP.

PA (UYST-) UNIV STELLENBOSCH.

XX Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;

PI WPI; 2002-154920/20.

XX New polynucleotides encoding antigenic HIV Type C polypeptides, useful

PT in applications including DNA immunization or generation of packaging

PT cell lines, particularly in gene therapy -

PS Claim 7; Fig 22; 233pp; English.

XX The present invention describes expression cassettes comprising a

CC polynucleotide sequence encoding a polypeptide comprising immunogenic

CC HIV type C polypeptides. The expression cassettes comprise any of the

CC HIV type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or

CC Nef (I). (I) have immunostimulant activity and can be used in gene

CC therapy. The HIV type C polynucleotides are useful in applications

CC including DNA immunisation, generation of packaging coli lines, and

CC production of HIV type C proteins. The polynucleotides are particularly

CC useful in gene therapy and DNA immunisation applications. ABL39942 to

CC ABL40054 and ABB06204 to ABB06215 represent sequences used in the

CC exemplification of the present invention.

XX Sequence 1494 BP; 325 A; 557 C; 455 G; 157 T; 0 other;

SQ

Query Match: 86.0%; Score 1297.6; DB 24; Length 1494;

Best Local Similarity 92.7%; Pred. No. 4.1e-153;

Matches 1402; Conservative 0; Mismatches 89; Indels 21; Gaps 3;

Qy 1 ATGGGCGCCCGCGCCAGCATCTCTGCGCGGCGGAGAGCTGGACAACTGGGAGAGATCCGC 60

Db 1 ATGGGCGCCCGCGCCAGCATCTCTGAGCGGCGGCGGCAAGCTGGACAACTGGGAGCCATCCGC 60

Qy 61 CTGGCGCCCGCGGCGGCAAGAGCACTACATGCTGAAGCACCTGTTGGGCCACCGCGGAG 120

Db 61 CTGGCGCCCGCGGCGGCAAGAGCACTACATGCTGAAGCACCTGTTGGGCCACCGCGGAG 120

QY 121 CTGGAGGGCTTCGCCCTGACCCCGGCGCTCTGGAGACCGCCGAGGCGTGCACAGCAGATC 180
DB 121 CTGGAGGGCTTCGCCCTGACCCCGGCGCTCTGGAGACCGCCGAGGCGTGCACAGCAGATC 180
QY 181 ATGAAGCAGCTGCAGCCCGGCGCTGCAGACCGGACCGGAGAGCTGCCGAGCCTGACACAC 240
DB 181 ATGAAGCAGCTGCAGCCCGGCGCTGCAGACCGGACCGGAGAGCTGCCGAGCCTGACACAC 240
QY 241 ACCGTGGCCACCTGTACTGCGTGCACCGCGGCGATCGAGGTCCGCGACACCAAGAGGCC 300
DB 241 ACCGTGGCCACCTGTACTGCGTGCACCAAGGCGATCGAGGTCCGCGACACCAAGAGGCC 300
QY 301 CTGGACAAGATCAGGAGGAGGAGACACAAGTCCGAGCAGAGACCCGAGAGGCCAAGGAG 360
DB 301 CTGGACAAGATCAGGAGGAGGAGGAGACACAAGTCCGAGCAGAGGCCGAGAGGCCAAGGAG 360
QY 361 GCGGACGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGSCCAGATGGTG 420
DB 361 GCGGACGGAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGSCCAGATGGTG 420
QY 421 CACGAGGCCATCAGCCCGCGCCACCTGAAAGCGCTGGTGAAGGTGATCGAGGAGAGGCC 480
DB 421 CACGAGGCCATCAGCCCGCGCCACCTGAAAGCGCTGGTGAAGGTGATCGAGGAGAGGCC 480
QY 481 TTCAGCCCGGAGGTGATCCCGATGTTCACCGCGCTGAGCGAGGCGGCCACCCCGAGGAG 540
DB 481 TTCAGCCCGGAGGAGATCCCGATGTTCACCGCGCTGAGCGAGGCGGCCACCCCGAGGAG 540
QY 541 CTGAACACCATGCTGAACACCGCTGGCGGCGGCGGACAGGCGGCGCATGCAGATCTGAAGG 600
DB 541 CTGAACACCATGCTGAACACCGCTGGCGGCGGCGGACAGGCGGCGCATGCAGATCTGAAGG 600
QY 601 ACCATCAAGCAGGAGGCGCGGAGTGGGACCGCTGCACCGCTGCAGGCGGCGCGCGCTG 660
DB 601 ACCATCAAGCAGGAGGCGCGGAGTGGGACCGCGCTGCACCGCGCTGCAGGCGGCGCGCTG 660
QY 661 GCGCCCGGCGAGATGCGGCGGACCGCGGCGGCGGAGACATGCGGCGGCGGCGGAGCAGG 720
DB 661 GCGCCCGGCGAGATGCGGCGGCGGCGGCGGCGGAGACATGCGGCGGCGGCGGAGCAGG 720
QY 721 CAGGAGCAGATCGCTGGATGACGAGCAGCAACCGCGCGTGGCGGCGGCGGAGATCTACAG 780
DB 721 CAGGAGCAGATCGCTGGATGACGAGCAGCAACCGCGCGTGGCGGCGGCGGAGATCTACAG 780
QY 781 CGGTGGATCATCTGGCGCTGAACAAGATCGTGGGAGTACAGCGCGGCGGAGATCTCTG 840
DB 781 CGGTGGATCATCTGGCGCTGAACAAGATCGTGGGAGTACAGCGCGGCGGAGATCTCTG 840
QY 841 GACATCGCGCGGCGGCGGCGGAGGCGCTTCGCGGAGTACGTGGACCGCTTCTTCAAGAC 900
DB 841 GACATCAAGCAGGCGGCGGCGGAGGCGCTTCGCGGAGTACGTGGACCGCTTCTTCAAGAC 900
QY 901 CTGGCGCGGAGCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
DB 901 CTGGCGCGGAGCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
QY 961 CAGAACGCCAACCGGAGTGCAGACCATCTCTGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1020
DB 961 CAGAACGCCAACCGGAGTGCAGACCATCTCTGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1020
QY 1021 GAGGAGATGATGACCGCTGCCAGGCGGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
DB 1021 GAGGAGATGATGACCGCTGCCAGGCGGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
QY 1081 GCGAGGCGATGAGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140
DB 1081 GCGAGGCGATGAGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140
QY 1141 GGCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200
DB 1138 GGCGAACCGCATCATCAAGTGGCTCAACTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1197

QY 1201 TGCGCGCGCGCGCGCAAGAAGGGCTGCTGGAGTGGCGCAAGGAGGGCGCAGAGTGAAG 1260
DB 1198 TGCGCGCGCGCGCGCAAGAAGGGCTGCTGGAGTGGCGCAAGGAGGGCGCAGAGTGAAG 1257
QY 1261 GACTGCACCGAGCGCGCAGGCCAACTTCCTGGCAAGATCTGGCCCGCAGCCCAAGGGCGCG 1320
DB 1256 GACTGCACCGAGCGCGCAGGCCAACTTCCTGGCAAGATCTGGCCCGCAGCCCAAGGGCGCG 1317
QY 1321 CCGCGCAACTTCCTGCAGAACCGCAGCGGAGC---CGCGCGCGCGCGCAGCGTGGCCACCGCG 1377
DB 1318 CCGCGCAACTTCCTGCAGAACCGCGCGCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1377
QY 1378 CCGCGCGCGGAGAGCTTCGCTTCGAGGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1437
DB 1378 CCGCGCGCGGAGAGCTTCGCTTCGAGGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1437
QY 1438 GACCGCGAGCGCTACCGCGAGCGCGCTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1497
DB 1438 -----CGCGAGCGCGCTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1492
QY 1498 CTGAGCGCGAGTAA 1509
DB 1483 CTGAGCGCGAGTAA 1494
RESULT 6
ABL40020
ID ABL40020 standard; DNA; 1491 BP.
XX ABL40020;
AC ABL40020;
XX ABL40020;
DT 15-MAY-2002 (first entry)
XX
DE Synthetic Gag polynucleotide sequence SEQ ID NO:99.
XX
KW Human immunodeficiency virus type C; antigenic HIV type C protein; immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef; immunostimulant; gene therapy; gene; ds.
KW
XX Human immunodeficiency virus type C.
OS
OS Synthetic.
XX
PN WC200204493-A2.
XX
XX 17-JAN-2002.
XX
PF 05-JUL-2001; 2001WO-US21241.
XX
PR 05-JUL-2000; 2000US-0610313.
XX
PA (CHIR) CHIRON CORP.
PA (UYST-) UNIV STELLENBOSCH.
XX
PI Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;
XX
XX WPI; 2002-154920/20.
XX
XX New polynucleotides encoding antigenic HIV Type C polypeptides, useful in applications including DNA immunization or generation of packaging cell lines, particularly in gene therapy .
PS Claim 8; Fig 70; 233pp; English.
XX
CC The present invention describes expression cassettes comprising a polynucleotide sequence encoding a polypeptide comprising immunogenic HIV type C polypeptides. The expression cassettes comprise any of the HIV type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or Nef (i). (i) have immunostimulant activity and can be used in gene therapy. The HIV type C polynucleotides are useful in applications including DNA immunisation, generation of packaging cell lines, and production of HIV Type C proteins. The polynucleotides are particularly useful in gene therapy and DNA immunisation applications. ABL39942 to ABL40054 and ABB06204 to ABB06215 represent sequences used in the

PA (CHIR ; CHIRON CORP.
XX Barnett S, Zur Megele C;
XX WPI; 2000-452401/39.
DR P-PSDB; AAY96943.
XX
PT Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV
PT Env polypeptide and the polypeptide useful for immunizing a mammal
PT especially human against HIV
XX
PS Claim 2; Page 92-93; 113pp; English.
XX
CC Expression cassettes comprising a polynucleotide encoding antigenic
CC type C human immunodeficiency virus (HIV) Gag or Env polypeptides are
CC useful in DNA immunization, generation of packaging cell lines and
CC production of Gag- and/or Env-containing proteins. Synthetic Env and Gag
CC expression cassettes exhibit increased potency for induction of
CC cytotoxic T-lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1
CC self-assemble into non-infectious virus-like particles which are used as
CC a matrix for the proper presentation of an antigen entrapped or
CC associated to the immune system of the host.
XX
SQ Sequence 1479 BP; 325 A; 529 C; 463 G; 162 T; 0 other;

Query Match 84.6%; Score 1276.8; DB 21; Length 1479;
Best Local Similarity 92.5%; Pred. No. 1.6e-150;
Matches 1399; Conservative 0; Mismatches 77; Indels 36; Gaps 4;

QY 1 ATGGGGCGGGCGCCAGCATCTGCGGGCGAGAGCTGACAAAGTGGAGAGATCCGC 60
DB 1 ATGGGGCGGGCGCCAGCATCTGCGGGCGAGCAAGCTGGAGGCGATCCGC 60

QY 61 CTGCGCCCGGGCGGCAAGCACTACATGCTGAAGCACTGGTGTGGCGCAGCGCGAG 120
DB 61 CTGCGCCCGGGCGGCAAGCACTACATGATGAAGCACTGGTGTGGCGCAGCGCGAG 120

QY 121 CTGAGGGGCTTCCGCTGAACCCCGGCTGCTGGAGACCGCGAGGCTGCAAGCAGATC 180
DB 121 CTGAGAGAGTTCCGCTGAACCCCGGCTGCTGGAGACCGCGAGGCTGCAAGCAGATC 180

QY 181 ATGAAGCAGCTGCAGCCCGCGCTGCAGACCGGACCGAGGAGCTGCGAGGCTGTACAG 240
DB 181 ATCGCGCAGCTGCAGCCCGCGCTGCAGACCGGCGGAGGAGCTGAAGAGCTGTACAG 240

QY 241 ACCGTGGCCACCGCTGTACTGGGCGACCGCGAGCTGAGGTCGCGACACCAAGGAGGC 300
DB 241 ACCGTGGCCACCGCTGTACTGGGCGACCGAGAGATCGAGTCCGCGACACCAAGGAGGC 300

QY 301 CTGGACACAGATCGAGGAGGAGAGAGCAAGTCCGAGCAGAGACCCAGAGCCCAAGAG 360
DB 301 CTGGACACAGATCGAGGAGGAGAGAGCAAGTCCGAGCAGAGATCCAGAGCCCAAGAG 360

QY 361 GCGGAC---GGCAAGGTGAGCGAGAACTACCCCATCGTGCAGAACTGCGAGGCGCAGATG 417
DB 361 GCGGACAAAGGCAAGGTGAGCGAGAACTACCCCATCGTGCAGAACTGCGAGGCGCAGATG 420

QY 418 GTGACACAGGCGCATCAGCCCGCGACCCCTGAACCGCTGGTGAAGGTGATCGAGGAGAG 477
DB 421 GTGACACAGGCGCATCAGCCCGCGACCCCTGAACCGCTGGTGAAGGTGATCGAGGAGAG 480

QY 478 GCCTTCAGCCCGGAGGTGATCCCATGTTACCGCCCTGAGCGAGGCGGCCACCCCGCAG 537
DB 481 GCCTTCAGCCCGGAGGTGATCCCATGTTACCGCCCTGAGCGAGGCGGCCACCCCGCAG 540

QY 538 GACCTGAACACAGATGTTGAACACCGTGGCGGCGCACCCAGCGCGCCATGCAATGCTGAG 597
DB 541 GACCTGAACACAGATGTTGAACACCGTGGCGGCGCACCCAGCGCGCCATGCAATGCTGAG 600

QY 598 GACACCATCAACGAGGAGGCGCGGAGTGGAGCGGCTGCACCCCGTGCAGCGCGCGCC 657
DB 601 GACACCATCAACGAGGAGGCGCGGAGTGGAGCGGCTGCACCCCGTGCAGCGCGCGCC 660

QY 658 GTGGCCCGCGCCAGATGCGGAGACCCCGCGGAGGAGCATGCGCGGCGCCACAGCACC 717
DB 661 ATCGCCCGCGCCAGATGCGGAGACCCCGCGGAGGAGCATGCGCGGCGCCACAGCACC 720

QY 718 CTGAGGAGGAGATCGCTGGAAGACCAACACCGCGCGTGGCGGAGACATCTAC 777
DB 721 CTGAGGAGGAGATCGCTGGAAGACCAACACCGCGCGTGGCGGAGACATCTAC 780

QY 778 AAGCGGTGATCATCTGGGCTGAACAGATGCTGCGGATGATACAGCCCGGTGAGCATC 837
DB 781 AAGCGGTGATCATCTGGGCTGAACAGATGCTGCGGATGATACAGCCCGGTGAGCATC 840

QY 838 CTGGACATCCGCGAGGCGCCAAAGGAGCGCTCCGCGACTACGTGGAGCGCTTCTTCAAG 897
DB 841 CTGGACATCAAGCAGGCGCCAAAGGAGCGCTCCGCGACTACGTGGAGCGCTTCTTCAAG 900

QY 898 ACCCTGCGCGGAGCAGGCGCCAGGAGCGTGAAGAACTGGAAGACCGACCCCTGCTG 957
DB 901 ACCCTGCGCGGAGCAGGCGCCAGGAGCGTGAAGAACTGGAAGACCGACCCCTGCTG 960

QY 958 GTGCAGAACGCCAACCCCGACTGCAAGACCATCTGCGCGCTCTCGGCGCGCGCGCAC 1017
DB 961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCTGCGCGCTCTCGGCGCGCGCGCAC 1020

QY 1018 CTGAGGAGATGATGACCGCTGCCAGGCGGTGGCGCGCGCGCGCACAGGCGCGGTG 1077
DB 1021 CTGAGGAGATGATGACCGCTGCCAGGCGGTGGCGCGCGCGCGCACAGGCGCGGTG 1080

QY 1078 CTGGCCGAGCGATGAGCCAGGCGCAACAGCGTGAACATCATGATGCAAGAGCAACTTC 1137
DB 1081 CTGGCCGAGCGATGAGCCAGGCGCAACAGCGTGAACATCATGATGCAAGAGCAACTTC 1137

QY 1138 AAGGCGCGCGCGCGCAAGTCAAGTGTTCAACTGCGGCAAGGAGGCGCGCACATCGCC 1197
DB 1138 AAGGCGCGCGCGCGCAAGTCAAGTGTTCAACTGCGGCAAGGAGGCGCGCACATCGCC 1197

QY 1198 AACTGCGCGCGCGCGCAAGAGGCGTGTGGAAGTGGCGCAAGGAGGCGCGCACAGATG 1257
DB 1198 AACTGCGCGCGCGCGCAAGAGGCGTGTGGAAGTGGCGCAAGGAGGCGCGCACAGATG 1257

QY 1258 AAGGACTGCACCGAGCGCGCGCAACTTCTTGGGCAAGATCTGGCCAGCCACAGGCG 1317
DB 1258 AAGGACTGCACCGAGCGCGCGCAACTTCTTGGGCAAGATCTGGCCAGCCACAGGCG 1317

QY 1318 CGCGCGCGCACTTCTTGCAGAACCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1377
DB 1318 CGCGCGCGCACTTCTTGCAGAACCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1359

QY 1378 CCGCGCGCGGAGAGCTTCCGCTTGGAGGAGACCAACCGCGCGCGCGCGCGCGCGCG 1437
DB 1380 CCGCGCGCGGAGAGCTTCCGCTTGGAGGAGACCAACCGCGCGCGCGCGCGCGCGCG 1419

QY 1438 GACCGCGAGCGCTACCGCGAGCGCGCTGACCGCGCGCTGCGGAGCGCTGTTGCGGAG 1497
DB 1420 G-----ACCGCGAGCGCTGACCGCGCGCTGAGAGCGCTGTTGCGGAGCGCGCG 1467

QY 1498 CTGAGCCAGTAA 1509
DB 1468 CTGAGCCAGTAA 1479

RESULT 10
AAL44548
ID AAL44548 standard; DNA; 1479 BP.
XX AAL44548;
XX
D7 08-NOV-2002 (first entry)
XX
DE HIV-1 p55gag polypeptide coding sequence 1.
XX
KW HIV; ds; vaccine; gene; immune response; microparticle;
KW adsorbent surface; poly(alpha-hydroxy acid); polyhydroxy butyric acid;

KW polycaprolactone; polyorthoester; polycyanacrylate; detergent;
KW submicron emulsion; viral infection; bacterial infection;
XX parasitic infection; HIV-1 p55gag polypeptide.

OS Human immunodeficiency virus type 1.

XX WO200226209-A2.

PN 04-APR-2002.

XX 28-SEP-2001; 2001WO-US30540.

XX 28-SEP-2000; 2000US-236105P.

PR 30-AUG-2001; 2001US-315905P.

XX (CHIR) CHIRON CORP.

XX O'hagan D, Otten G, Donnelly JJ, Polo JM, Barnett S, Singh M;

PI Ulmer J, Dubensky JW;

XX WPI; 2002-519084/55.

XX A microparticle to which a biologically active macromolecule is
PT adsorbed, for use as a vaccine composition to treat viral, bacterial or
PT parasitic infections, comprises a polymer microparticle, a detergent
PT and a submicron emulsion -

PS Claim 72; Fig 1; 100pp; English.

XX The invention relates to a method of raising an immune response in a host
CC animal. The method of the invention comprises administering a
CC microparticle that has an adsorbent surface to which a first biologically
CC active macromolecule (e.g. a nucleic acid) has been adsorbed. The
CC microparticle comprises a polymer microparticle of poly(alpha-hydroxy
CC acid), a polyhydroxy butyric acid, a polycaprolactone, a polyorthoester,
CC a polycyanacrylate, a detergent, and submicron emulsion. The method/
CC microparticle of the invention is useful for immunising a host animal
CC against viral, bacterial or parasitic infections. The present DNA
CC sequence encodes a HIV-1 p55gag polypeptide.

XX Sequence 1479 BP: 325 A: 529 C: 463 G: 162 T: 0 other;

Query Match 84.6%; Score 1276.8; DB 24; Length 1479;

Best Local Similarity 92.5%; Pred. No. 1.6e-150;

Matches 1399; Conservative 0; Mismatches 77; Indels 36; Gaps 4;

QY 1 ATGGGGCGCGCGCCAGCATCTCGCGCGGAGAACCTGGACAAAGTGGGAGAGATCCGC 60

DB 1 ATGGGGCGCGCGCGCCAGCATCTCGCGCGGCGCAAGCTGGACGCTGGGAGCGCATCCGC 60

QY 61 CTGGCGCGCGCGCGCGCAAGAGCACTACATGCTGAAGCAACCTGGTGGCGCGCGGAG 120

DB 61 CTGGCGCGCGCGCGCGCAAGAGTGTACATGATGAAGCAACCTGGTGGCGCGCGGAG 120

QY 121 CTGGAGGGCTTCGCCCTGAACCCCGCGCTGCTGGAGACCGCGCGAGGCTGCAAGCAGATC 180

DB 121 CTGGAGAGATTGCCCCGAGAACCCCGCGCTGCTGGAGACCGCGCGAGGCTGCAAGCAGATC 180

QY 181 ATGAAGCAGCTGCAGCGCGCGCTGCAGACCGCGCGAGGAGCTGCGCAGCGCTGCACAA 240

DB 181 ATCCGCCAGCTGCACCGCGCGCTGCAGACCGCGCGAGGAGCTGAAGAGCGCTGTCAAG 240

QY 241 ACCGTGCGCACCCCTGTACTCGCTGCACCGCGCGCATCGAGGTCCGCGACACCAAGAG 300

DB 241 ACCGTGCGCACCCCTGTACTCGCTGCACCGCGCGCATCGAGGTCCGCGACACCAAGAG 300

QY 301 CTGGACAGATCGAGGAGGAGCAGAACAAAGTCCAGCAGAAAGACCCAGCAGCGCGAGAG 360

DB 301 CTGGACAGATCGAGGAGGAGCAGAACAAAGTCCAGCAGCAGAAATCCAGCAGCGCGAG 360

QY 361 GCGGAC---GGCAAGGTGAGCCAGAACTACCCCATCTGTGCAGAACTGTGCGGCGCGAGATG 420

DB 361 GCGGACAAAGGCTGAGCCAGAACTACCCCATCTGTGCAGAACTGTGCGGCGCGAGATG 420

QY 418 GTGCACAGGGCCATCAGCCCCCGCCACCTGAACGCTGGTGAAGGTGATCGAGGAGAAG 477
DB 421 GTGCACAGGGCCATCAGCCCCCGCCACCTGAACGCTGGTGAAGGTGATCGAGGAGAAG 480
QY 478 GCCTTCAGCGCGCGAGGTGATCCCATGTTTACCAGCCCTGAGCGAGGGCGCCACCCCGAG 537
DB 481 GCCTTCAGCGCGCGAGGTGATCCCATGTTTACCAGCCCTGAGCGAGGGCGCCACCCCGAG 540
QY 538 GACCTGAACACGATGTTGAACACCGTGGCGCGGCACACGAGCCGCCCATGCAGATGCTGAAG 597
DB 541 GACCTGAACACGATGTTGAACACCGTGGCGCGGCACACGAGCCGCCCATGCAGATGCTGAAG 600
QY 598 GACACCATCAACGAGGAGGGCGCGCGAGTGGGACCCGCTGCACCCGCTGCAGGCGGGCGCC 657
DB 601 GACACCATCAACGAGGAGGGCGCGCGAGTGGGACCCGCTGCACCCGCTGCAGGCGGGCGCC 660
QY 658 GTGGCCCCCGCGCGAGATCGCGGAGCCCCCGCGCGAGCGACATGGCGGGCGCCACCGACCC 717
DB 661 ATGGCCCCCGCGCGAGATCGCGGAGCCCCCGCGCGAGCGACATGGCGGGCGCCACCGACCC 720
QY 718 CTGCAGGAGCAGATCGCCTGGATGACACGCAACCCCGCTGCCCGTGGCGGACATCTAC 777
DB 721 CTGCAGGAGCAGATCGCCTGGATGACACGCAACCCCGCTGCCCGTGGCGGACATCTAC 780
QY 778 AAGCGGTGGATCATCTGGCGCTGAACAAGATCGTCGGGATGTACAGCCCCCGTGGAGCATC 837
DB 781 AAGCGGTGGATCATCTGGCGCTGAACAAGATCGTCGGGATGTACAGCCCCCGTGGAGCATC 840
QY 838 CTGGACATCCCGCGAGGGCCCCAAGGAGCCCTTCCGGGACTACGTGGACCGCTTCTTCAAG 897
DB 841 CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGGGACTACGTGGACCGCTTCTTCAAG 900
QY 898 ACCCTGCGCGCGAGCGGCCACCCAGGACGCTGAACAACCTGGATGACCGGAGACCTGCTG 957
DB 901 ACCCTGCGCGCGAGCAGACACCCAGGAGGTGAACAACCTGGATGACCGGAGACACCTGCTG 960
QY 958 GTGCAGAACGCCAACCCCGACTGCAAGACCATCTCTGGCGCGCTCTCTGGCGCCCGCGCCACC 1017
DB 961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCTCTGGCGCGCTCTCTGGCGCCCGCGCCACC 1020
QY 1018 CTGGAGGAGATGATGACCGCTTCCAGGGCGTGGCGGGCCCCGGCCACAGGCGCGCGTG 1077
DB 1021 CTGGAGGAGATGATGACCGCTTCCAGGGCGTGGCGGGCCCCGGCCACAGGCGCGCGTG 1080
QY 1078 CTGGCGGAGCGATGAGCCAGGCCACAGCGTGAACATCATGATGCAGAAAGAGCAACTTC 1137
DB 1081 CTGGCGGAGCGGATGAGCCAGGCCACACACC---AGCGTGTATGATGCAGAAAGAGCAACTTC 1137
QY 1138 AAGGGCCCCCGCGCGCAAGTCTCAAGTCTCAACTCTGGCGCAAGGAGGCGCCACATCGCCCGC 1197
DB 1138 AAGGGCCCCCGCGCGCAAGTCTCAAGTCTCAACTCTGGCGCAAGGAGGCGCCACATCGCCCGC 1197
QY 1198 AACTGCGCGCGCGCGCGCAAGAGGGTCTGGAAGTGGCGCAAGGAGGGCGCCACCATG 1257
DB 1198 AACTGCGCGCGCGCGCGCAAGAGGGTCTGGAAGTGGCGCAAGGAGGGCGCCACCATG 1257
QY 1258 AAGGACTGCACCGGAGCGCGCGCAACTTCTCTGGGCAAGATCTGGCGCGAGCCACAGGGCC 1317
DB 1258 AAGGACTGCACCGGAGCGCGCGCAACTTCTCTGGGCAAGATCTGGCGCGAGCCACAGGGCC 1317
QY 1318 CGCCCCCGCAACTTCTCTGAGAACCGCGAGCGCCCGCGCGCCCGCCACCGTGCACCGCC 1377
DB 1318 CGCCCCCGCAACTTCTCTGAGAGCGG-----CCCCGAGCCCGCCCGCC 1359
QY 1378 CCCCCCGCGGAGAGCTTCGGCTTCGAGGAGACCCACCCCGCGCGCGCGCGCGCGCGCGAG 1437
DB 1360 CCCCCCGCGGAGAGCTTCGGCTTCGAGGAGACCCACCCCGCGCGCGCGCGCGCGAGAGCAAG 1419
QY 1438 GACCGGAGCCCTACCGGAGCGCCCTGACCGCGCGCTGCGCAGCCCTGTGCGGCGAGCGCGCC 1497
DB 1420 G-----ACCGGAGAGCCCTGACCGAGCGCTGGAAGAGCGCTGTTCGGCAACCGACCC 1467

QY 1498 CTGAGCCAGTAA 1509
Db 1468 CTGAGCCAGTAA 1479

RESULT 11
ID ABL39953
XX ABL39953 standard; DNA; 1479 BP.
AC ABL39953;
DI 15-MAY-2002 (first entry)
XX Synthetic Gag polynucleotide sequence SEQ ID NO:3.
DE Human immunodeficiency virus type C; antigenic HIV type C protein;
KW immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
KW immunostimulant; gene therapy; gene; ds.
XX Human immunodeficiency virus type C.
OS Synthetic.
OS WO20020493-A2.
PN 17-JAN-2002.
PD 05-JUL-2001; 2001WO-US21241.
XX 05-JUL-2000; 2000US-0610313.
PR (CHIR) CHIRON CORP.
PA (UYST-) UNIV STELLENBOSCH.
XX Zur Megede J. Barnett SW. Engelbrecht S. Van Rensburg EJ.
PI WPI: 2002-154920/20.
DR
XX New polynucleotides encoding antigenic HIV Type C polypeptides, useful
PT in applications including DNA immunization or generation of packaging
PT cell lines, particularly in gene therapy .
XX Example 1; Fig 1; 233pp; English.
XX The present invention describes expression cassettes comprising a
CC polynucleotide sequence encoding a polypeptide comprising immunogenic
CC HIV type C polypeptides. The expression cassettes comprise any of the
CC HIV type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or
CC Nef (I). (I) have immunostimulant activity and can be used in gene
CC therapy. The HIV type C polynucleotides are useful in applications
CC including DNA immunisation, generation of packaging cell lines, and
CC production of HIV Type C proteins. The polynucleotides are particularly
CC useful in gene therapy and DNA immunisation applications. ABL39942 to
CC ABL40054 and ABB06204 to ABB06215 represent sequences used in the
CC exemplification of the present invention.
XX Sequence 1479 BP; 325 A; 529 C; 463 G; 162 T; 0 other;
SQ

Query Match 84.6%; Score 1276.8; DB 24; Length 1479;
Best Local Similarity 92.5%; Pred. No. 1.6e-150;
Matches 1399; Conservative 0; Mismatches 77; Indels 36; Gaps 4;

QY 1 ATGGGGCCCCGCGCAGCATCTGCGCGGCGAGAAAGCTGGACAAAGTGGGAGAAATCCGC 60
Db 1 ATGGGGCCCCGCGCAGCATCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 60

QY 61 CTGCGCCCCCGCGGCGAAGAGTGTACATGATGAAGCACCTTGTGTGGGCGAGCCGAG 120
Db 61 CTGCGCCCCCGCGGCGAAGAGTGTACATGATGAAGCACCTTGTGTGGGCGAGCCGAG 120

QY 121 CTGAGGGGCTTCGCCCTGAAACCCCGGCTGTGTGAGAGACCCGAGGGCTGCAAGCAGATC 180
Db 121 CTGAGAGAGTTCGCCCTGAACCCCGGCTGTGTGAGAGACCCGAGGGCTGCAAGCAGATC 180

QY 181 ATGAAGCAGCTGCAGCCCGCCCTGCAGACGCGCACCGAGGAGCTGCGAGCCCTGTACAAC 240
Db 181 ATCCGCCAGCTGCACCCCGCCCTGCAGACGCGCACCGAGGAGCTGAAGAGCCCTGTACAAC 240

QY 241 ACCGTGGCCACCCCTGTACTGGGTGCACGCGCGGCTCGAGGTCCGCGGACACCAAGGAGGCC 300
Db 241 ACCGTGGCCACCCCTGTACTGGGTGCACGAGAAAGATCGAUGTCCGCGGACACCAAGGAGGCC 300

QY 301 CTGGACAAGATCGAGGAGGAGAGCAAGTCCAGCAGAGACCCAGCAGGCAAGGAG 350
Db 301 CTGGACAAGATCGAGGAGGAGAGCAAGTCCAGCAGAGATCCAGAGGCGCGAGGCC 350

QY 361 GCGGAC--GGCAAGGTGAGCCAGAACTACCCCATCTGTCAGAACTGCGAGGGCCAGATG 417
Db 361 GCGGACAAGGCAAGGTGAGCCAGAACTACCCCATCTGTCAGAACTGCGAGGGCCAGATG 420

QY 418 CTGCACCAAGGATCATGAGCCCGCCAGCCCTGAACGCTTGGTGAAGTGTGAGGAGAG 477
Db 421 GTGCACCAAGGATCATGAGCCCGCCAGCCCTGAACGCTTGGTGAAGTGTGAGGAGAG 480

QY 478 GCTTTCAGCCCGAGGTGATCCCCCATGTTACCCGCTGAGCGAGGCGCCACCCCGAG 537
Db 481 GCTTTCAGCCCGAGGTGATCCCCCATGTTACCCGCTGAGCGAGGCGCCACCCCGAG 540

QY 538 GACCTGAACACGATGTTGAACACCGTGGCGGCGCCACAGCCGCTGAGGCGCGCCG 597
Db 541 GACCTGAACACGATGTTGAACACCGTGGCGGCGCCACAGCCGCTGAGGCGCGCCG 600

QY 598 GACACCATCAACGAGGAGGCGCGGAGTGGGACCGCTGAGACCCGCTGAGGCGCGCCG 657
Db 601 GACACCATCAACGAGGAGGCGCGGAGTGGGACCGCTGAGACCCGCTGAGGCGCGCCG 660

QY 658 GTGGCCCCCGGCGCAGATGCGGACCCCGCGGCGAGCGACATCGCGCGCCACCAAGCACC 717
Db 661 ATCGCCCCCGGCGCAGATGCGGAGCCCGCGGCGAGCGACATCGCGCGCCACCAAGCACC 720

QY 718 CTGCAGGAGCAGATCGCTGTGATGACACCGTGGCGGCGCCACAGCCGCTGAGGCGACATC 777
Db 721 CTGCAGGAGCAGATCGCTGTGATGACACCGTGGCGGCGCCACAGCCGCTGAGGCGACATC 780

QY 778 AACGGTGGATCATCTCTGGGCTGAACAAGATCGTGGGATGTACAGCCCGCTGAGCATC 837
Db 781 AACGGTGGATCATCTCTGGGCTGAACAAGATCGTGGGATGTACAGCCCGCTGAGCATC 840

QY 838 CTGGACATCCGCGAGGCGCCCAAGGAGCCCTTCCGCGACTAGCTGGAGCCGCTTCTTCAAG 897
Db 841 CTGGACATCAACGAGGCGCCCAAGGAGCCCTTCCGCGACTAGCTGGAGCCGCTTCTTCAAG 900

QY 898 ACCCTGGCGCGCGAGCAGGCGCCACCCAGGACGCTGAAGAAGTGGATGACCGAGACCCCTGCTG 957
Db 901 ACCCTGGCGCGCGAGCAGGACCCAGGAGTGAAGAAGTGGATGACCGAGACCCCTGCTG 960

QY 958 GTGCAGAACGCGCAACCCCGACTGCAAGACCATCTGCGGCTCTCGGCGCCCGCGGCGCAC 1017
Db 961 GTGCAGAACGCGCAACCCCGACTGCAAGACCATCTGCGGCTCTCGGCGCCCGCGGCGCAC 1020

QY 1018 CTGGAGGAGATGATGACCGCTGCCAGGCGCTGGCGCGCCCGCGCCACAGGCGCCCGCTG 1077
Db 1021 CTGGAGGAGATGATGACCGCTGCCAGGCGCTGGCGCGCCCGCGCCACAGGCGCCCGCTG 1080

QY 1078 CTGGCGGAGGCGATGAGCCAGCCCAACAGCGTGAACATCATGATGCAAGAGAGCAACTTC 1137
Db 1081 CTGGCGGAGGCGATGAGCCAGCCCAACACC--AGCGTGAATGTCAGAGAGCAACTTC 1137

QY 1138 AAGGGCCCCCGGCGCAACGTCGAAGTGTCTCAACTGCGGCAAGGAGGCGCCACATCGCCAG 1197
Db 1138 AAGGGCCCCCGGCGCAACGTCGAAGTGTCTCAACTGCGGCAAGGAGGCGCCACATCGCCAG 1197

QY 1198 AACTGCGCGCGCGCGCGCGAGAGGGCTGTCTGGAAGTGGCGCAAGGAGGCGCCACAGATG 1257
Db 1198 AACTGCGCGCGCGCGCGCGAGAGGGCTGTCTGGAAGTGGCGCAAGGAGGCGCCACAGATG 1257

QY 1258 AAGGACTGCACCGAGCGCCAGGCCAACTTCTCTGGGCAAGATCTGGCCCGCCACCAAGGCG 1317

Db 967 GTGCAGAACGCCAACCCGACTGCAAGACCATCCGCGGCTCTCGGCGCCGCGCCAGC 1025
Qy 1018 CTGGAGGAGATGATGACGCGCTGCCAGGCGGTGGCGGCCCGGCCGACAAAGCGCGGTG 1077
Db 1027 CTGGAGGAGATGATGACCGGCTGCCAGGCGGTGGCGGCCCGGCCGACAAAGCGCGGTG 1086
Qy 1078 CTGGCCGAGGCGATGAGCCAGGCCCAACAGCGTGAACATCATGATGCAGAAGAGCACTTC 1137
Db 1087 CTGGCCGAGGCGATGAGCCAGGCCCAACAGC--AGCGTATGATGCAGAAGAGCACTTC 1143
Qy 1138 AAGGGCCCCCGCGCAACGTCAGTCTCAACTGCGCAAGGAGGGGCGACATCCGCAAG 1197
Db 1144 AAGGGCCCCCGCGCATCGTCAAGTCTCAACTGCGCAAGGAGGGGCGACATCGCGCGC 1203
Qy 1198 AACTGCGCGCGCCCGCGCAAGAGGGCTGCTGGAAGTGGCGCAAGGAGGGCCACCATG 1257
Db 1204 AACTGCGCGCGCCCGCGCAAGAGGGCTGCTGGAAGTGGCGCAAGGAGGGCCACCATG 1263
Qy 1258 AAGGACTGCACCGAGCGCCAGGCCCAACTCTCTGGGCAAGATCTGGCCAGCCACAGGGC 1317
Db 1264 AAGGACTGCACCGAGCGCCAGGCCCAACTCTCTGGGCAAGATCTGGCCAGCCACAGGGC 1323
Qy 1318 CGCCCCGCAACTTCCTGCAGAACCGCGAGCGCGCGCCCGCCCGCCCGCCCGCCCGCC 1377
Db 1324 CGCCCCGCAACTTCCTGCAGAGCGG-----CGCGAGCGCCCGCCCGCCCGCCCGCC 1365
Qy 1378 CCCCCCGCGAGAGCTTCGCTTCGAGGAGAACCCCGCGCCCGCCCGCCCGCCCGCCCGCC 1437
Db 1366 CCCCCCGCGAGAGCTTCGCTTCGAGGAGAACCCCGCGCCCGCCCGCCCGCCCGCCCGCC 1425
Qy 1438 GACCGCGAGCGCTACCGCGAGCGCCCTGACCGCGCTGCGGAGCGCTGTTGCGAGCGCGCC 1497
Db 1426 G-----ACCGGAGACCCCTGACCGAGCGCTGAGAGCGCTGTGCGGCAAGCGCGC 1473
Qy 1498 CTGAGCCAGTAA 1509
Db 1474 CTGAGCCAGAAA 1485

RESULT 13
ACA03522
ID ACA03522 standard; DNA; 4715 BP.

xx ACA03522;
AC ACA03522;
XX 22-MAY-2003 (first entry)
DX Synthetic DNA encoding immunogenic HIV peptide #5.
DE Immunogenic HIV polypeptide; human immunodeficiency virus; HIV;
KW vaccine; gene therapy; packaging cell line; humoral immune response;
KW cellular immune response; gene delivery vector; DNA immunisation;
KW ds.
XX Synthetic.
OS
XX
XX WO2003004657-A1.
PN
XX
PD 16-JAN-2003.
XX
XX 05-JUL-2002; 2002WO-US21421.
PF
XX
PR 05-JUL-2001; 2001US-303192P.
PR 31-AUG-2001; 2001US-316860P.
PR 16-JAN-2002; 2002US-349728P.
PR 16-JAN-2002; 2002US-349793P.
PR 16-JAN-2002; 2002US-349871P.
XX
XX (CHIR) CHIRON CORP.
PA
PI Zur Megede J, Barnett SW, Lian Y;
XX

DR WPI: 2003-221602/21.
XX New synthetic polynucleotides encoding antigenic HIV type B and/or type
PI C polypeptides, useful as immunogenic compositions or vaccines for
PI generating humoral or cellular immune responses against HIV in a
PI subject, especially humans -
XX
PS Example 1; Fig 10; 262pp; English.
XX The invention describes a synthetic polynucleotide encoding 2 or more
CC immunogenic HIV polypeptides, where at least 2 of the polypeptides are
CC derived from different HIV subtypes. The polynucleotide is useful for
CC immunisation, generation of packaging cell lines, or production of HIV
CC polypeptides. The polynucleotide and its encoded proteins are useful as
CC immunogenic compositions or vaccines for generating humoral or cellular
CC immune responses against HIV in a subject, or for inducing neutralising
CC antibodies against HIV. The gene delivery vector comprising the
CC polynucleotide is also useful for DNA immunisation of, or for
CC generating an immune response (e.g. a humoral or cellular immune
CC response) in a subject such as a mammal, particularly a human. This
CC sequence encodes a human immunodeficiency virus immunogenic peptide.
XX
SQ Sequence 4716 BP; 1041 A; 1684 C; 1463 G; 528 T; 0 other;

Query Match: 84.5%; Score 1275.2; DB 25; Length 4716;
Rest Local Similarity 92.5%; Pred. No. 2e-150;
Matches 1398; Conservative 0; Mismatches 78; Indels 36; Gaps 4;

Qy 1 ATGGGGCGCCCGCCAGCATCTCTGCGGGCGGAGAGCTGGACAAGTGGGAGAGATCCGC 60
Db 7 ATGGGGCGCCCGCCAGCATCTCTGCGGGCGGAGAGCTGGACAAGTGGGAGAGATCCGC 66
Qy 61 CTGCGCCCCCGGGCGCAAGAAAGTCTACATCTGTAAGCACCTGGTGTGCGCCAGCGCGAG 120
Db 67 CTGCGCCCCCGGGCGCAAGAAAGTCTACATGATGAAGCACCTGGTGTGCGCCAGCGCGAG 126
Qy 121 CTGAGGGGCTTCGCGCTTGAACCCCGCGCTCTGCTGGAGACCGCGAGGCTGCAAGCAGATC 180
Db 127 CTGGAGAAGTTCGCGCTTGAACCCCGCGCTCTGCTGGAGACCGCGAGGCTGCAAGCAGATC 186
Qy 181 ATGAAGCAGCTGCAGCCCGCGCTGCAGACCGCGACCGAGGAGCTGCGCAGCGCTGTACAC 240
Db 187 ATCGCCAGCTGCAGCCCGCGCTGCAGACCGCGAGGAGCTGAGAGCGCTGTACAC 246
Qy 241 ACCGTGGCCACCTGTACTGCTGCAGCCCGCGCTGCAGCTGCGGACACCAAGGAGGCC 300
Db 247 ACCGTGGCCACCTGTACTGCTGCAGGAGAGATCGAGGTTCGCGACCAAGGAGGCC 306
Qy 301 CTGCACAAGATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
Db 307 CTGCACAAGATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 366
Qy 361 GCGGAC--GGCAAGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 417
Db 367 GCGGACAGGGCAAGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 426
Qy 418 GTGCACCGAGGCGATCAGCCCGCGCTGCAAGCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 477
Db 427 GTGCACCGAGGCGATCAGCCCGCGCTGCAAGCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 486
Qy 478 GCCTTCAGCCCGCGAGGTGATCCCATGTTACCGCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 537
Db 487 GCCTTCAGCCCGCGAGGTGATCCCATGTTACCGCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 546
Qy 538 GACCTGAACACGATGTTGAACACCGTGGGCGGCGACCAAGGCGCGCATGCGAGATGCTGAAG 597
Db 547 GACCTGAACACGATGTTGAACACCGTGGGCGGCGACCAAGGCGCGCATGCGAGATGCTGAAG 606
Qy 598 GACACCATCAACGAGGAGGCGCGGAGTGGGAGCGCTGCACCCCGCTGCAGGCGCGCGGCC 657
Db 607 GACACCATCAACGAGGAGGCGCGGAGTGGGAGCGCTGCACCCCGCTGCAGGCGCGCGGCC 666
Qy 658 GTGGCCCCCGCGCAGATGCGCGGAGCCCCCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 717

```
Db 667 ATCGCCCGCGGACATCGCGAGCCCGCGGAGCGGACATCGCGGACCGACCG 726
QY 718 CTGAGGAGCAGATCGGCTGGATGACCAAGCAACCGCGGTGGGTGACATCTAG 777
Db 727 CTGAGGAGCAGATCGGCTGGATGACCAAGCAACCGCGGTGGGTGACATCTAG 786
QY 778 AAGCGGTGATCATCGTGGGCTGAACAAGATCGTGGGATGTACAGCCCGTGAAGATC 837
Db 787 AAGCGGTGATCATCGTGGGCTGAACAAGATCGTGGGATGTACAGCCCGTGAAGATC 846
QY 838 CTGACATCCCGCAGGCGCCCAAGGAGCCCTTCGCGACTAGCTGGAACCGCTCTTCAAG 897
Db 847 CTGACATCAAGCAGGCGCCCAAGGAGCCCTTCGCGACTAGCTGGAACCGCTCTTCAAG 906
QY 898 ACCCTGCGCGCGAGCAGGCGCCCAAGGAGCGTGAAGAAGTGAATGACCGAGACCGTGTG 957
Db 907 ACCCTGCGCGCGAGCAGGCGCCCAAGGAGCGTGAAGAAGTGAATGACCGAGACCGTGTG 966
QY 958 GTGCAAGACGCAACCGCGACTGCAAGACCATCTCTGGGCTGTGGCCCGCGCGCAAC 1017
Db 967 GTGCAAGACGCAACCGCGACTGCAAGACCATCTCTGGGCTGTGGCCCGCGCGCAAC 1026
QY 1018 CTGAGGAGATGATGACCGCGCTGCGAGGGCGTGGCGGCGCCGCGCAAGCGCGGG 1077
Db 1027 CTGAGGAGATGATGACCGCGCTGCGAGGGCGTGGCGGCGCCGCGCAAGCGCGGG 1086
QY 1078 CTGCGCGAGGCGATGAGCCAGCGCAACAGCGTGAACATCATGTATGACAGAGCAACTTC 1137
Db 1087 CTGCGCGAGGCGATGAGCCAGCGCAACAGCGTGAACATCATGTATGACAGAGCAACTTC 1146
QY 1138 AAGGCGCCCGCGGCAACGCTCACTGCTTCAAGTGGCGCAAGGAGGCGCCAGCGCAAG 1197
Db 1144 AAGGCGCCCGCGGCAACGCTCACTGCTTCAAGTGGCGCAAGGAGGCGCCAGCGCAAG 1203
QY 1198 AACTGCGCGCGCGCGCGCAAGAGCGCTGTGGAAGTGGCGCAAGGAGGCGCGCAAGATG 1257
Db 1204 AACTGCGCGCGCGCGCGCAAGAGCGCTGTGGAAGTGGCGCAAGGAGGCGCGCAAGATG 1263
QY 1258 AAGGACTGACCGAGCGCGCGAGCGCAACTTCTGGCGCAAGATCTGGCGCGCGCAAGG 1317
Db 1264 AAGGACTGACCGAGCGCGCGAGCGCAACTTCTGGCGCAAGATCTGGCGCGCGCAAGG 1323
QY 1318 CGCCCGCGCAACTTCTGCAAGACCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1377
Db 1324 CGCCCGCGCAACTTCTGCAAGAGCGCG-----CGCCGAGCGCGCGCGCGCGCG 1365
QY 1378 CCGCGCGCGGAGGCTTCGCGTTGAGGAGACCGCGCGCGCGCGCGCGCGCGCGCGCG 1437
Db 1366 CCGCGCGCGGAGGCTTCGCGTTGAGGAGACCGCGCGCGCGCGCGCGCGCGCGCGCG 1425
QY 1438 GACCGGAGCGCGTACCGCGGAGCGCGCTGACCGCGCTGCGGAGCGCTGTTGGCAGCGCG 1497
Db 1426 G-----ACCGCGAGCGCGTACCGAGCGCTGAAGAGCGCTGTTGGCAGCGCGCG 1473
QY 1498 CTGAGCCAGTAA 1509
Db 1474 CTGAGCCAGAAA 1485
```

```
RESULT 14
ACA03524
ID ACA03524 standard: DNA: 2742 BP.
XX
AC ACA03524;
XX
DT 22-MAY-2003 (first entry)
XX
DE Synthetic DNA encoding immunogenic HIV peptide #7.
XX
KW Immunogenic HIV polypeptide; human immunodeficiency virus; HIV;
KW vaccine; gene therapy; packaging cell line; humoral immune response;
KW cellular immune response; gene delivery vector; DNA immunisation;
```

```
KW ds.
XX Synthetic.
OS WC2003004657-A1.
XX
XX 16-JAN-2003.
XX
XX 05-JUL-2002; 2002WO-US21421.
XX
XX 05-JUL-2001; 2001US-303192P.
XX 31-AUG-2001; 2001US-316360P.
XX 16-JAN-2002; 2002US-349728P.
XX 16-JAN-2002; 2002US-349793P.
XX 16-JAN-2002; 2002US-349871P.
XX (CHIR ) CHIRON CORP.
XX
XX Zur Mevede J, Barnett SW, Lian Y;
XX WPI; 2003-221602/21.
XX
XX New synthetic polynucleotides encoding antigenic HIV type B and/or type
XX C polypeptides, useful as immunogenic compositions or vaccines for
XX generating humoral or cellular immune responses against HIV in a
XX subject, especially humans -
XX
XX Example 1: Fig 12; 262bp; English.
XX
XX The invention describes a synthetic polynucleotide encoding 2 or more
XX immunogenic HIV polypeptides, where at least 2 of the polypeptides are
XX derived from different HIV subtypes. The polynucleotide is useful for
XX immunisation, generation of packaging cell lines, or production of HIV
XX polypeptides. The polynucleotide and its encoded proteins are useful as
XX immunogenic compositions or vaccines for generating humoral or cellular
XX immune responses against HIV in a subject, or for inducing neutralising
XX antibodies against HIV. The gene delivery vector comprising the
XX polynucleotide is also useful for DNA immunisation of, or for
XX generating an immune response (e.g. a humoral or cellular immune
XX response) in, a subject such as a mammal, particularly a human. This
XX sequence encodes a human immunodeficiency virus immunogenic peptide.
XX
XX Sequence 2742 BP; 578 A; 1020 C; 859 G; 285 T; 0 other;
XX
XX Query Match 84.4%; Score 1273.6; DB 25; Length 2742;
XX Best Local Similarity 92.4%; Pred. No. 3.5e-150;
XX Matches 1397; Conservative 0; Mismatches 79; Indels 36; Gaps 4;
QY 1 ATGGGCGCGCGCGCGCGCATCTCTGCGCGGCGGAGAGCTGGACACTGGGAGAGATCCGC 60
Db 7 ATGGGCGCGCGCGCGCGCATCTCTGCGCGGCGGAGAGCTGGACACTGGGAGAGATCCGC 66
QY 61 CTGGGCGCGCGCGCGCGCAAGAGCACTACATGCTGAGACACCTGTGTGGCGCGCGCGAG 120
Db 67 CTGGGCGCGCGCGCGCGCAAGAGTGTACATGATGAGACACCTGTGTGGCGCGCGCGAG 126
QY 121 CTGGAGGCGCTTCGCCCTGAACCGCGCGCTGCTGGAGACCGCGCGAGGCTGCRAGCAGATC 180
Db 127 CTGGAGAGTTCGCCCTGAACCGCGCGCTGCTGGAGACCGCGCGAGGCTGCRAGCAGATC 186
QY 181 ATGAAGCAGCTGCAAGCGCGCGCTGAGAGCGCGCGCGCGAGGAGCTGCGCAGCGCTGACAC 240
Db 187 ATCGGCGAGCTGCACCGCGCGCTGAGAGCGCGCGCGCGAGGAGCTGAAGAGCGCTTTCAC 246
QY 241 ACCGTGGCGCGCGCTTACTGCGTGCACCGCGCGCTCGAGGTCGCGCGACACCGAGGCGCC 300
Db 247 ACCGTGGCGCGCGCTTACTGCGTGCACCGAGAGATCGAGGTGCGCGACACCGAGGCGCC 306
QY 301 CTGACAAAGATCGAGGAGGAGCAGACAAAGTCCCGAGCAGAGACCGCGCGCGCGAGGAG 360
Db 307 CTGACAAAGATCGAGGAGGAGCAGACAAAGTCCCGAGCAGAGACCGCGCGCGCGAGGCG 366
QY 361 GCGGAC---GGCAGGCTGAGCGCGCGAGCACTACCGCATCGTGCAGACCTGCGAGCGCGCAGATG 417
```

Db 367 GCGACAAGGCAAGGTGAGCCAGAACTACCCCATCTGTGAGAACCTGTAGGGCCAGATG 426
QY 418 GTGCACCAAGCCATCAGCCCGCCGACCCCTGAAGCCCTGGTGAAGTGTGAGAGAGAG 477
Db 427 GTGCACCAAGCCATCAGCCCGCCGACCCCTGAAGCCCTGGTGAAGTGTGAGAGAGAG 486
QY 478 GCCTTCAGCCCGAGGTGATCCCATGTTACCGCCCTGAGCGAGGGGACCCCCAG 537
Db 487 GCCTTCAGCCCGAGGTGATCCCATGTTACCGCCCTGAGCGAGGGGACCCCCAG 546
QY 538 GACCTGAACACCATGTGTAACACCGTGGGGGCCACCAAGCGCGCCATGCAATGCTGAAG 597
Db 547 GACCTGAACACCATGTTGAACACCGTGGGGGCCACCAAGCGCGCCATGCAATGCTGAAG 606
QY 598 GACACCATCAACAGGAGGCGCGCGAGTGGGACCGCCCTGCACCCCGTGCAGGCGCGCC 657
Db 607 GACACCATCAACAGGAGGCGCGCGAGTGGGACCGCGCTGCACCCCGTGCAGGCGCGCC 666
QY 658 GTGGCCCCCGGCAGATGCGCGACCCCGCGCGAGCGACATCGCCGGCGCCACCAAGCAC 717
Db 667 ATCGCCCCCGGCAGATGCGCGAGCCCGCGCGAGCGACATCGCCGGCACCAAGCAC 726
QY 718 CTGAGGAGCAGATCGCTGGATGACCAAGACCCCGCGCGAGCGACATCGCTGAGCATCTAC 777
Db 727 CTGAGGAGCAGATCGCTGGATGACCAAGACCCCGCGCGAGCGACATCGCTGAGCATCTAC 786
QY 778 AAGCGGTGGATCTCTGGGCGTGAACAAGATCGTGGGATGTACAGCCCGCTGAGCATCT 837
Db 787 AAGCGGTGGATCTCTGGGCGTGAACAAGATCGTGGGATGTACAGCCCGCTGAGCATCT 846
QY 838 CTGACATCTCGCCAGGCGCCCAAGAGCCCTTCGCGACTAGCTGGACCGCTCTTCAAG 897
Db 847 CTGACATCAAGCAGGCGCCCAAGAGCCCTTCGCGACTAGCTGGACCGCTCTTCAAG 906
QY 898 ACCCTGGCGCGAGCGCCACCGAGGACGTGAAGAAGTGAATGACCGAGACCGCTGCTG 957
Db 907 ACCCTGGCGCGAGCAGCGCCACCGAGGAGGTGAAGAAGTGAATGACCGAGACCGCTGCTG 967
QY 958 GTGCAGAACGCCAACCCGACTGCAAGACCATCTGCGGCTCGCGCCCGCGCGGAC 1017
Db 967 GTGCAGAACGCCAACCCGACTGCAAGACCATCTGCGGCTCGCGCCCGCGCGGAC 1026
QY 1018 CTGAGGAGATGATGACCGCTGCGAGGCGGTGGCGGCCCGCGCAAGCGCGCGTG 1077
Db 1027 CTGAGGAGATGATGACCGCTGCGAGGCGGTGGCGGCCCGCGCAAGCGCGCGTG 1086
QY 1078 CTGCGCGAGGATGAGCGCAGGCCAACAGCGGTGAACATCATGTGCAAGAGCAACTTC 1137
Db 1087 CTGCGCGAGGATGAGCGCAGGCCAACACCC---AGCGTGTATGTGCAAGAGCAACTTC 1143
QY 1138 AAGGCCCCCGCGCAACGTCAAGTCTTCAACTGCGGCAAGGAGGCGCCACATCGCCAAAG 1197
Db 1144 AAGGCCCCCGCGCAACGTCTCAAGTCTTCAACTGCGGCAAGGAGGCGCCACATCGCCCGC 1203
QY 1198 AACTGCGCGCCCGCGCAAGAGGCTGCTGGAAGTGGCGCAAGGAGGCGCCACAGATG 1257
Db 1204 AACTGCGCGCCCGCGCAAGAGGCTGCTGGAAGTGGCGCAAGGAGGCGCCACAGATG 1263
QY 1258 AAGGACTGCACCGAGCGCGCAAGTCTTCCGGAAGATCTGGCCCGCGCAAGGCGCC 1317
Db 1264 AAGGACTGCACCGAGCGCGCAAGTCTTCCGGAAGATCTGGCCCGCGCAAGGCGCC 1323
QY 1318 CGCCCGCGCAACTTCTCTGCAAGACCGCAGCGCGCGCCCGCCACCGCTGCGCACCGCC 1377
Db 1324 CGCCCGCGCAACTTCTCTGCAAGCGC-----CGCCGAGCGCCACCGCC 1365
QY 1378 CGCCCGCGCGAGAGCTTCGCGCTCGAGGAGACCCACCGCCCGCCCAAGCAGGAGCCCAAG 1437
Db 1366 CGCCCGCGCGAGAGCTTCGCGCTTCGAGGAGACCCACCGCCCGCGCAAGCAGGAGCAAG 1425
QY 1438 GACCGGAGCCCTACCGCGAGCCCTGACCGCCCTGCGCAGCCCTGTCGCGAGCGCGCC 1497

Db 1426 G-----ACCGCGAGACCTGACCAGCCTGAAGAGCCCTGCGCAACGACCCC 1473
QY 1498 CTGAGCCAGTAA 1509
Db 1474 CTGAGCCAAAGAA 1485
RESULT 15
ACA03521
ID ACA03521 standard; DNA; 5145 BP.
XX
AC ACA03521;
XX
DT 22-MAY-2003 (first entry)
XX
DE Synthetic DNA encoding immunogenic HIV peptide #4.
XX
KW immunogenic HIV polypeptide; human immunodeficiency virus; HIV;
KW vaccine; gene therapy; packaging cell line; humoral immune response;
KW cellular immune response; gene delivery vector; DNA immunisation;
KW ds.
XX
OS Synthetic.
XX
PN WC2003004657-A1.
XX
ID 16-JAN-2003.
XX
PE 05-JUL-2002; 2002WO-US21421.
XX
PR 05-JUL-2001; 2001US-303192P.
PR 31-AUG-2001; 2001US-316860P.
PR 15-JAN-2002; 2002US-349728P.
PR 16-JAN-2002; 2002US-349793P.
PR 16-JAN-2002; 2002US-349871P.
XX
PA (CHIR) CHIRON CORP.
XX
PI Zur Megede J, Barnett SW, Lian Y;
XX
WP1: 2003-221602/21.
XX
PT New synthetic polynucleotides encoding antigenic HIV type B and/or type
P: C polypeptides, useful as immunogenic compositions or vaccines for
PI generating humoral or cellular immune responses against HIV in a
PT subject, especially humans -
XX
PS Example 1; Fig 9; 262pp; English.
XX
CC The invention describes a synthetic polynucleotide encoding 2 or more
CC immunogenic HIV polypeptides, where at least 2 of the polypeptides are
CC derived from different HIV subtypes. The polynucleotide is useful for
CC immunisation, generation of packaging cell lines, or production of HIV
CC polypeptides. The polynucleotide and its encoded proteins are useful as
CC immunogenic compositions or vaccines for generating humoral or cellular
CC immune responses against HIV in a subject, or for inducing neutralising
CC antibodies against HIV. The gene delivery vector comprising the
CC polynucleotide is also useful for DNA immunisation of, or for
CC generating an immune response (e.g. a humoral or cellular immune
CC response) in, a subject such as a mammal, particularly a human. This
CC sequence encodes a human immunodeficiency virus immunogenic peptide.
XX
SQ Sequence 5145 BP; 1130 A; 1842 C; 1599 G; 574 T; 0 other;

Query Match 84.4%; Score 1273.6; DB 25; Length 5145;
Best Local Similarity 92.4%; Pred. No. 3.1e-150;
Matches 1397; Conservative 0; Mismatches 79; Indels 36; Gaps 4;
QY 1 ATGGGCGCGCGCCAGCATCTCTGCGCGCGGAGAGCTGGACAAAGTGGGAGAGATCCGC 60
Db 7 ATGGGCGCGCGCCAGCATCTCTGCGCGCGGCGCAAGCTGGACGCGCTGGGAGCGCATCCGC 66
QY 61 CTGCGCCCGCGCGGCGCAAGACACTACATGCTGAAGACCTGGTGGGGCCAGCCCGGAG 120

Db 67 CTGCGCCCGCGGCAAGAGTGTACATGATGAAGCACCTGGTGTGGCCAGCCGCGAG 126
QY 121 CTGGAGGGTTGCGCCCTGAACCCCGCCTGCTGGAGACGGGAGGGGTGCAAGCAGATG 190
Db 127 CTGGAGAAGTTCGCCCTGAACCCCGCCTGCTGGAGACCGAGCGGCTGCAAGCAGATG 186
QY 181 ATGAAGCAJGCGAGCCCGCTGCGAGACCGGACCGAGSAGCTGCGCAGCTGTACAAC 240
Db 187 ATCCGCCAGCTGCAACCCCGCCTGCGAGACCGGACCGAGGAGCTGAAGAGCTGTTCAC 246
QY 241 ACCGTGGCCACCTGTACTGCTGTCACCGCCCGCATCGAGTCCCGGACACCAAGGAGGCC 300
Db 247 ACCGTGGCCACCTGTACTGCTGTCACCGCCCGCATCGAGTCCCGGACACCAAGGAGGCC 306
QY 301 CTGGACAAGATCGAGGAGGAGGAGACAACAAGTCCACGACAGAACCCAGCAGGCCCAAGAG 350
Db 307 CTGGACAAGATCGAGGAGGAGGAGACAACAAGTGCACGACAGAAATCCAGCAGCCGAGGCC 356
QY 361 GCCGAC--GGCAAGGTGAGCCACGACCCCTGAAGCCCTGGTGACAACTGCGAGGCCAGATG 417
Db 367 GCGGACAAGGGCAAGGTGAGCCAGAACTACCCCATCTGTCAGAACTGTCAGAGCCGAGT 426
QY 418 GTGCACCAAGGCCATCAGCCCGGACCCCTGAAGCCCTGGTGACAACTGCGAGGCCAGATG 477
Db 427 GTGCACCAAGGCCATCAGCCCGGACCCCTGAAGCCCTGGTGACAACTGTCAGAGGCCAG 486
QY 478 GCGTTGAGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 537
Db 487 GCGTTGAGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 546
QY 538 GACCTGAACACGATGTGAACACCGTGGGCGGACACCGGCGCCCACTGACAGTGTGAAG 597
Db 547 GACCTGAACACGATGTGAACACCGTGGGCGGACACCGGCGCCCACTGACAGTGTGAAG 606
QY 598 GACACCATCAACGAGGAGGCGCGGAGTGGGAGCGCCTGCACCCCTGCAGGCCGCGCCCG 657
Db 607 GACACCATCAACGAGGAGGCGCGGAGTGGGAGCGCCTGCACCCCTGCAGGCCGCGCCCG 666
QY 658 GTGGCCCGCGGCGGAGATCGCGGACCCCGGCGGAGCGACATCGCGCGGCGGCGGAGCACC 717
Db 667 ATCGCCCGCGGCGGAGATCGCGGAGCCCGCGGCGGAGCGACATCGCGCGGCGGCGGAGCACC 726
QY 718 CTGCAGGAGCAGATCGCCTGGATGACCAAGCAAGCCCGCGCTGCGCGTGGGCGACATCTAC 777
Db 727 CTGCAGGAGCAGATCGCCTGGATGACCAAGCAAGCCCGCGCTGCGCGTGGGCGACATCTAC 786
QY 778 AAGCGGTGATCATCTCTGGCCCTGAACAAGATGCTGGGATGTACAGCCCGCTGAGCATC 837
Db 787 AAGCGGTGATCATCTCTGGCCCTGAACAAGATGCTGGGATGTACAGCCCGCTGAGCATC 846
QY 838 CTGCACATCGCGGAGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 897
Db 847 CTGCACATCAAGCAGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 906
QY 898 ACCCTGCGCGCGGAGCAGCGCCACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 957
Db 907 ACCCTGCGCGCGGAGCAGCGCCACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 966
QY 958 GTGCAGAACCGCAACCCCGCTGCAAGACCATCTGCGCGCTGTGGCCCGGCGGAGGAGGAG 1017
Db 967 GTGCAGAACCGCAACCCCGCTGCAAGACCATCTGCGCGCTGTGGCCCGGCGGAGGAGGAG 1026
QY 1018 CTGGAGGAGATGATGACCGCTGCGAGGCGGTGGGCGGCGCCCGGCGGAGGAGGAGGAGGAG 1077
Db 1027 CTGGAGGAGATGATGACCGCTGCGAGGCGGTGGGCGGCGCCCGGCGGAGGAGGAGGAGGAG 1086
QY 1078 CTGGCGGAGGATGAGCCAGGCGCAACAGCGGTGAACATCATGATGACAGAGAGCAACTTC 1137
Db 1087 CTGGCGGAGGATGAGCCAGGCGCAACACC---ACGCTGATGATGACAGAGAGCAACTTC 1143
QY 1138 AAGGCGCCCGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1197

Db 1144 AAGGCGCCCGCGGCGGATCGTCAAGTGTCTCAACTCGGCGCAAGGAGGCGGCACTCGCCCGC 1203
QY 1198 AACTGCGCGCCCGCGGCAAGAGAGGGCTGCTGGAAGTGGGCAAGGAGGCGGCACTGATG 1257
Db 1204 AACTGCGCGCCCGCGGCAAGAGAGGGCTGCTGGAAGTGGGCAAGGAGGCGGCACTGATG 1263
QY 1258 AAGGACTGACCGAGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1317
Db 1264 AAGGACTGACCGAGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1323
QY 1318 CGCCCGGCAACTTCTGTCAGAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1377
Db 1324 CGCCCGGCAACTTCTGTCAGAGGCGG-----CCCCGAGCCACCGGCC 1365
QY 1378 CCCCCCGCGAGAGGCTTCGCTTCGAGGAGACCCCGCCCGCCCGCCCGGAGGAGGAGGAGGAG 1437
Db 1386 CCCCCCGCGGAGAGGCTTCGCTTCGAGGAGACCCCGCCCGCCCGGAGGAGGAGGAGGAGGAG 1425
QY 1438 GACCGGAGCCCTACCGGAGCCCTGACCGCCCTGACCGCCCTGCGGAGGAGGAGGAGGAGGAG 1497
Db 1426 G-----ACCGGAGAGCCCTGACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1473
QY 1498 CTGAGCCAGTAA 1509
Db 1474 CTGAGCCCAAGAA 1485

Search completed: September 22, 2003. 17:48:06
Job time : 446.317 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: September 22, 2003, 17:26:03 ; Search time 5707.74 Seconds
(without alignments)
10815.605 Million cell updates/sec

Title: US-09-475-704A-4
Perfect score: 1509
Sequence: 1 atggggcccgccagcal.....gcggcccccctgagccagtaa 1509

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :		GenEmbl:*	
1:	gb_ba:*	21	1117.8
2:	gb_htg:*	22	1117.8
3:	gb_in:*	23	1103.8
4:	gb_om:*	24	1103.8
5:	gb_ov:*	25	1103.8
6:	gb_pat:*	26	1100
7:	gb_ph:*	27	1097.8
8:	gb_pl:*	28	1097.8
9:	gb_pr:*	29	1097.8
10:	gb_ro:*	30	1084.8
11:	gb_sts:*	31	1082.6
12:	gb_sy:*	32	1075.6
13:	gb_un:*	33	1072.2
14:	gb_vt:*	34	1071.8
15:	em_ba:*	35	1070.2
16:	em_fun:*	36	1064.2
17:	em_hum:*	37	981
18:	em_in:*	38	971.4
19:	em_mu:*	39	930.2
20:	em_om:*	40	930.2
21:	em_or:*	41	929.4
22:	em_ov:*	42	929.4
23:	em_pat:*	43	929.4
		44	925.4
		45	925.4

AF202465	Synthetic	12	1847	74.1
AX427928	Sequence	6	1847	74.1
AX427921	Sequence	6	9179	73.1
AX427932	Sequence	6	9189	73.1
AX427931	Sequence	6	9792	73.1
AX427933	Sequence	6	9169	72.9
AX427935	Sequence	6	9167	72.8
AX427934	Sequence	6	9782	72.8
AX427938	Sequence	6	9783	72.8
AX306428	Sequence	6	9785	71.9
AX306428	Sequence	6	4343	71.7
AF287352	Synthetic	12	4352	71.3
AX427937	Sequence	6	9407	71.1
AX306429	Sequence	6	4341	71.0
AX306430	Sequence	6	3981	70.9
AF287353	Synthetic	12	4353	70.5
AX457086	Sequence	6	1905	65.0
BD057031	Synthetic	6	1532	64.4
AX035453	Sequence	6	4307	61.6
AX056836	Sequence	6	4307	61.6
AR302563	Sequence	6	4307	61.6
AX019132	Sequence	6	4307	61.6
BD130468	Antiviral	6	4307	61.6
AX283597	Sequence	6	4307	61.3
AX283607	Sequence	6	4307	61.3

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1509	100.0	1509	6	AX455888	AX455888 Sequence
2	1507.4	99.9	1509	6	AX468544	AX468544 Sequence
3	1504.2	99.7	1509	6	AX468548	AX468548 Sequence
4	1489.8	98.7	1509	6	AX455905	AX455905 Sequence
5	1297.6	86.0	1494	6	AX455935	AX455935 Sequence
6	1294.8	85.8	1491	6	AX455983	AX455983 Sequence
7	1276.8	84.6	1479	6	AX455887	AX455887 Sequence
8	1276.8	84.6	1479	6	AX468543	AX468543 Sequence
9	1271.2	84.2	1479	6	AX468547	AX468547 Sequence
10	1260.8	83.6	1479	6	AX455904	AX455904 Sequence
11	1202.2	79.7	4288	6	AX149648	AX149648 Sequence
12	1201	79.6	1509	12	AF201927	AF201927 Synthetic
13	1186	78.6	9166	6	AX427930	AX427930 Sequence
14	1170.8	77.6	9788	6	AX427936	AX427936 Sequence
15	1145	75.9	1539	6	AX306431	AX306431 Sequence
16	1143.2	75.8	1548	12	AF287354	AF287354 Synthetic
17	1119	74.2	6438	6	AX427919	AX427919 Sequence
18	1119	74.2	9194	6	AX427925	AX427925 Sequence
19	1119	74.2	9194	6	AX427926	AX427926 Sequence
20	1119	74.2	12411	6	AX427927	AX427927 Sequence
21	1117.8	74.1	1847	12	AF202464	AF202464 Synthetic
22	1117.8	74.1	1847	12	AF202465	AF202465 Synthetic
23	1103.8	73.1	9170	6	AX427928	AX427928 Sequence
24	1103.8	73.1	9189	6	AX427921	AX427921 Sequence
25	1103.8	73.1	9792	6	AX427932	AX427932 Sequence
26	1100	72.9	9169	6	AX427931	AX427931 Sequence
27	1097.8	72.8	9167	6	AX427933	AX427933 Sequence
28	1097.8	72.8	9782	6	AX427935	AX427935 Sequence
29	1097.8	72.8	9783	6	AX427934	AX427934 Sequence
30	1084.8	71.9	9785	6	AX427938	AX427938 Sequence
31	1082.6	71.7	4343	6	AX306428	AX306428 Sequence
32	1075.8	71.3	4352	12	AF287352	AF287352 Synthetic
33	1072.2	71.1	9407	6	AX427937	AX427937 Sequence
34	1071.8	71.0	4341	6	AX306429	AX306429 Sequence
35	1070.2	70.9	3981	6	AX306430	AX306430 Sequence
36	1064.2	70.5	4353	12	AF287353	AF287353 Synthetic
37	981	65.0	1905	6	AX457086	AX457086 Sequence
38	971.4	64.4	1532	6	BD057031	BD057031 Synthetic
39	930.2	61.6	4307	6	AX035453	AX035453 Sequence
40	930.2	61.6	4307	6	AX056836	AX056836 Sequence
41	929.4	61.6	4307	6	AR302563	AR302563 Sequence
42	929.4	61.6	4307	6	AX019132	AX019132 Sequence
43	929.4	61.6	4307	6	BD130468	BD130468 Antiviral
44	925.4	61.3	4307	6	AX283597	AX283597 Sequence
45	925.4	61.3	4307	6	AX283607	AX283607 Sequence

ALIGNMENTS

RESULT 1
AX455888
LOCUS AX455888 1509 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 4 from Patent WO204493.
ACCESSION AX455888
VERSION AX455888.1 GI:21714881
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences
REFERENCE
1
AUTHORS zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.
TITLE Polynucleotides encoding antigenic hiv type c polypeptides,
polypeptides and uses thereof
JOURNAL Patent: WO 0204493-A 4 17-JAN-2002;

Db 121 CTGGAGGCGCTTCGCCCTGAACCCCGGCGCTCTGGAGACCGCGGAGGGCTGCAAGCAGATC 180

QY 181 ATGAAGCAGCTGCGAGCCCGCTGCAGACCGGCACCGAGAGCTGCGCAGCCTGTACAAC 240

Db 181 ATGAAGCAGCTGCGAGCCCGCTGCAGACCGGCACCGAGAGCTGCGCAGCCTGTACAAC 240

QY 241 ACCGTGGCCACCGTGTACTSCGTGCACCGCGGATCGAGGTTCGGCGGACACCAAGAGGCC 300

Db 241 ACCGTGGCCACCGTGTACTSCGTGCACCGCGGATCGAGGTTCGGCGGACACCAAGAGGCC 300

QY 301 CTGGACAAGATCGAGGAGGAGCAGAACAGTCCACGACAGAGACCCAGCAGGCCCAAGGAC 360

Db 301 CTGGACAAGATCGAGGAGGAGCAGAACAGTCCACGACAGAGACCCAGCAGGCCCAAGGAC 360

QY 361 GCCGACGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGCTG 420

Db 361 GCCGACGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGCTG 420

QY 421 CACGAGGCCATCAGCCCGCCCGACCCCTGAACGCTTGGGTGAAGGTGATCGAGGAGAGGCC 480

Db 421 CACGAGGCCATCAGCCCGCCCGACCCCTGAACGCTTGGGTGAAGGTGATCGAGGAGAGGCC 480

QY 481 TTCAGCCCGGAGGTGAICCCCATGTTTCACCGGCTGAGCGAGGGCGCCACCCCGCCAGGAC 540

Db 481 TTCAGCCCGGAGGTGAICCCCATGTTTCACCGGCTGAGCGAGGGCGCCACCCCGCCAGGAC 540

QY 541 CTGAACACGATGTTGAACACCGTGGCGGCCACCGAGCGCGCATGCAGATGCTGAAGGAC 600

Db 541 CTGAACACGATGTTGAACACCGTGGCGGCCACCGAGCGCGCATGCAGATGCTGAAGGAC 600

QY 601 ACCATCAACGAGGAGGCCCGGAGTGGGACCGCTGCACCCCGTGCAGGCGCGCCCGCTG 660

Db 601 ACCATCAACGAGGAGGCCCGGAGTGGGACCGCTGCACCCCGTGCAGGCGCGCCCGCTG 660

QY 661 GCCCGCGCGAGATGCGGACCCCGCGCGGACGACATCGCGCGCGCCACGACACCCCTG 720

Db 661 GCCCGCGCGAGATGCGGACCCCGCGCGGACGACATCGCGCGCGCCACGACACCCCTG 720

QY 721 CAGGAGCAGATCGCTGGATGACGAGCAACCCCGCTGCCCGTGGCGGACATCTACAAG 780

Db 721 CAGGAGCAGATCGCTGGATGACGAGCAACCCCGCTGCCCGTGGCGGACATCTACAAG 780

QY 781 CGGTGGATCACTCGGGCCTGAACAAGATCGTGGGATGTACAGCCCGCTGAGCATCCIG 840

Db 781 CGGTGGATCACTCGGGCCTGAACAAGATCGTGGGATGTACAGCCCGCTGAGCATCCIG 840

QY 841 GACATCGCGCGAGGCGCCCAAGGAGCCCTTCGCGGACTACGTGACCGCTTCTCAAGACC 900

Db 841 GACATCGCGCGAGGCGCCCAAGGAGCCCTTCGCGGACTACGTGACCGCTTCTCAAGACC 900

QY 901 CTGCGCGCGGAGCAGCGCCACCCAGGACGTTGAAGAACTGGATGACCGAGACCTGCTG 960

Db 901 CTGCGCGCGGAGCAGCGCCACCCAGGACGTTGAAGAACTGGATGACCGAGACCTGCTG 960

QY 961 CAGAACGCCAACCCCGACTGCAAGACCATCTGCGCGCTCTCGCGCCCGCGGCGCACCCCTG 1020

Db 961 CAGAACGCCAACCCCGACTGCAAGACCATCTGCGCGCTCTCGCGCCCGCGGCGCACCCCTG 1020

QY 1021 GAGGAGATGATGACCGCTGCGAGGGCGTGGCGCGCCCGCGGCGCACCAAGGCCCGCTG 1080

Db 1021 GAGGAGATGATGACCGCTGCGAGGGCGTGGCGCGCCCGCGGCGCACCAAGGCCCGCTG 1080

QY 1081 GCCGAGCGGATGAGCGAGGCCAACAGCGTGAACATCATGATGAGAGAGGACATCAAG 1140

Db 1081 GCCGAGCGGATGAGCGAGGCCAACAGCGTGAACATCATGATGAGAGAGGACATCAAG 1140

QY 1141 GGCCCGCGCGCAACGTTCAAGTCTCAACTGGGCGAAGGAGGCGCCACATCGCCCAAGNAC 1200

Db 1141 GGCCCGCGCGCAACGTTCAAGTCTCAACTGGGCGAAGGAGGCGCCACATCGCCCAAGNAC 1200

QY 1201 TGCCGCGCGCGCGCAAGAGGGCTGCTGGAAGTGGCGCAAGGAGGCGCCACAGATGAG 1260

Db 1201 TGCCGCGCGCGCGCAAGAGGGCTGCTGGAAGTGGCGCAAGGAGGCGCCACAGATGAG 1260

QY 1261 GACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCAGCCACAAAGGGCGC 1320

Db 1261 GACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCAGCCACAAAGGGCGC 1320

QY 1321 CCGGCAACTTCCTGCAGAACCGCAGCGAGCGCGCCCGCCCGCCACCGTGCACCGCCCG 1380

Db 1321 CCGGCAACTTCCTGCAGAACCGCAGCGAGCGCGCGCCCGCCCGCCACCGTGCACCGCCCG 1380

QY 1381 CCGGCGGAGAGCTTCGGCTTCGAGGAGACCAACCGCCCGCCCGCCCAAGCAGGAGCCCAAGGAC 1440

Db 1381 CCGGCGGAGAGCTTCGGCTTCGAGGAGACCAACCGCCCGCCCGCCCAAGCAGGAGCCCAAGGAC 1440

QY 1441 CGCGAGCCCTACCGGAGCGCCCTGACCGCCCTGCGCAGCCCTGTGCGCAGCGCCCGCTG 1500

Db 1441 CGCGAGCCCTACCGGAGCGCCCTGACCGCCCTGCGCAGCCCTGTGCGCAGCGCCCGCTG 1500

QY 1501 AGCCAGTAA 1509

Db 1501 AGCCAGTAA 1509

RESULT 4

AX455905

LOCUS AX455905 1509 bp DNA linear PAT 06-JUL-2002

DEFINITION Sequence 2i from Patent WO204493.

ACCESSION AX455905

VERSION AX455905.1 GI:21714897

KEYWORDS synthetic construct

SOURCE synthetic construct

ORGANISM artificial sequences.

REFERENCE 1

AUTHORS zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.

TITLE Polynucleotides encoding antigenic hiv type c polypeptides, polypeptides and uses thereof

JOURNAL Patent: WO 0204493-A 2i 17-JAN-2002;

CHIRON CORPORATION (US) : University of Stellenbosch (ZA)

FEATURES

location/Qualifiers

1..1509

/organism="synthetic construct"

/mol_type="genomic DNA"

/db_xref="taxon:32630"

/note="synthetic gag coding sequence of HIV strain AF110967"

BASE COUNT 321 a 559 c 471 g 158 t

ORIGIN

Query Match 98.7%; Score 1489.8; DB 6; Length 1509;

Best Local Similarity 99.2%; Pred. No. 8.7e-148;

Matches 1497; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 ATGGGGCCCCCGCCAGCATCCIGCGCGCGGAGAGACTGGACAAGTGGGAGAGATCCGC 60

Db 1 ATGGGGCCCCCGCCAGCATCCIGCGCGCGGAGAGACTGGACAAGTGGGAGAGATCCGC 60

QY 61 CTGCGCCCCCGCGCGCAAGAGCACTACATGCTGAAGCACCTGCTGGGCGCAGCGCGAG 120

Db 61 CTGCGCCCCCGCGCGCAAGAGCACTACATGCTGAAGCACCTGCTGGGCGCAGCGCGAG 120

QY 121 CTGGAGGGCTTCGCCCTGAACCCCGCCCTGCTGGAGACCCCGCAGGGCTGCAAGCAGATC 180

Db 121 CTGGAGGGCTTCGCCCTGAACCCCGCCCTGCTGGAGACCCCGCAGGGCTGCAAGCAGATC 180

QY 181 ATGAAGCAGCTGCAGCCCGCCCTGCAGACCCGGCACCGGAGGAGCTGCGCAGCCTGTACAAC 240

Db 181 ATGAAGCAGCTGCAGCCCGCCCTGCAGACCCGGCACCGGAGGAGCTGCGCAGCCTGTACAAC 240

QY 241 ACCGTGGCCACCCCTGTACTGCGTGCAGCGCGCGCATCGAGGTCCGCGACACCAAGAGGCC 300

Db 241 ACCGTGGCCACCCCTGTACTGCGTGCAGCGCGCGCATCGAGGTCCGCGACACCAAGAGGCC 300

QY 301 CTGGACAAGATCGAGGAGGAGCAGAACAGTCCACGACAGAGACCCAGCAGGCCCAAGGAG 360

481	TTCAACCCCGAGGAGATCCCCATGTTCCCGCCCTGAGCGAGGGCGCCACCCCGAGAC	542
541	CTGAACACGATGTTGAACACCGCTGGCGGCCACCAAGGCCCGCCATGCAGATGCTGAAGAC	600
541	CTGAACACCATGCTGAACACCGTGGCGGCCACCAAGGCCCGCCATGCAGATGCTGAAGAC	600
601	ACCATCAACGAGGAGGCCGCGAGTGGGACCGCCTGCACCCCGTGCAGGCGCGGCCCGT	660
601	ACCATCAACGAGGAGGCCGCGAGTGGGACCGCCTGCACCCCGTGCAGGCGCGGCCCGT	660
661	GCCCCGGCCAGATGCGGACCGCCCGCGGAGCGACATGCGCGCGGCCACAGCACCGCTG	720
661	GCCCCGGCCAGATGCGGACCGCCCGCGGAGCGACATGCGCGCGGCCACAGCACCGCTG	720
721	CAGGAGCAGATCGCCCTGGATGACCAAGACACCCCGCTGCGCGTGGGCGACATCTACAG	780
721	CAGGAGCAGATCGCCCTGGATGACCAAGACACCCCGCTGCGCGTGGGCGACATCTACAG	780
781	CGGTGATCATCTCTGGGCTGAACAAGATCGTGGGATGACAGCCCGCTGAGCATCTG	840
781	CGGTGATCATCTCTGGGCTGAACAAGATCGTGGGATGACAGCCCGCTGAGCATCTG	840
841	GACATCGCGCAGGCCCCCAAGGAGCCCTTCGCGGACTACCTGACCGCTTCTTCAGAGC	900
841	GACATCAACGAGGCCCCCAAGGAGCCCTTCGCGGACTACCTGACCGCTTCTTCAGAGC	900
901	CTGCGCGCCAGCAGGCCACCAAGGACGTGAAGAACTGATGACCGGAGCCCTGCTGGT	960
901	CTGCGCGCCAGCAGGCCACCAAGGACGTGAAGAACTGATGACCGGAGCCCTGCTGGT	960
961	CAGNACGCCAACCCCGACTGCAAGACCATCTCTGCGCGCTCTGGCGCCCGCCACCGCTG	1020
961	CAGAACGCCAACCCCGACTGCAAGACCATCTCTGCGCGCTCTGGCGCCCGCCACCGCTG	1020
1021	GAGGAGATGATGACCGCTGCGAGGGCGTGGCGCGCCCGCCACAGGCCCGCTGCTG	1080
1021	GAGGAGATGATGACCGCTGCGAGGGCGTGGCGCGCCCGCCACAGGCCCGCTGCTG	1080
1081	GCCGAGGCGATGAGCCAGGCCAACAGCGTGAACATCATGATGACAGAGAGCAACTTCAG	1140
1081	GCCGAGGCGATGAGCCAGGCCAACAGCGTGAACATCATGATGACAGAGAGCAACTTCAG	1140
1141	GGCCCCCGGCGCAACGTCAAGTGCTTCACTGCGGCAAGAGGGCCACATCGCCCAAGAC	1200
1141	GGCAGCAACCGCATCATCAAGTGCTTCACTGCGGCAAGAGGGCCACATCGCCCAAGAC	1200
1198	TGCGCGCGCCCGCCGCAAGAGGGCTGCTGGAAGTGGCGCCAGGAGGCCACAGATGAAG	1257
1198	TGCGCGCGCCCGCCGCAAGAGGGCTGCTGGAAGTGGCGCCAGGAGGCCACAGATGAAG	1257
1261	GACTGCACCGAGCGCCAGGCCAACTTCCCTGGCAAGATCTGCGCCCGCCACAGAGGGCGC	1320
1261	GACTGCACCGAGCGCCAGGCCAACTTCCCTGGCAAGATCTGCGCCCGCCACAGAGGGCGC	1320
1321	CCCGGCAACTTCTGCGAGAACCGCAGCGAGC--CCGCGCGCCCGCCACCGTGCACCGCGC	1377
1321	CCCGGCAACTTCTGCGAGAACCGCAGCGAGC--CCGCGCGCCCGCCACCGTGCACCGCGC	1377
1378	CCCCCGCGCAGAGCCTCCGCTTCGAGGAGACCAACCGCCCGCCCGCCAGAGAGCGCCAG	1437
1378	CCCCCGCGCAGAGCCTCCGCTTCGAGGAGACCAACCGCCCGCTGCCCGCAGAGAGAGGAG	1437
1438	GACCGCGAGCCCTACCGCGAGCCCTGACCGCCCTGCGGAGCCTGTTGCGAGCGCGCC	1497
1438	-----CGCGAGCCCTGACCGAGCCTGAAGAGCCTGTTGCGAGCGCGCC	1497
1498	CTGAGCCAGTAA	1509
1498	CTGAGCCAGTAA	1494

AX455983
LOCUS AX455983 1491 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 99 from Patent WO0204493.
ACCESSION AX455983
VERSION AX455983.1 GI:21714967
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1
AUTHORS zur Magede,J., Barnett,S.W., Engcibrecht,S. and van Rensburg,E.
TITLE Polynucleotides encoding antigenic hiv type c polypeptides,
polypeptides and uses thereof
JOURNAL Patent: WO 0204493-A 99 17-JAN-2002;
CHIRON CORPORATION (US); University of Stellenbosch (ZA)
FEATURES
source 1..1491
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="Gag_TV2_C_ZAopt"
BASE COUNT 319 a 557 c 457 g 158 t
ORIGIN
Query Match 85.8%; Score 1294.8; DB 6; Length 1491;
Rest_Score Similarity 92.7%; Pred. No. 2.6e-127;
Matches 1401; Conservative 0; Mismatches 87; Indels 24; Gaps 3;
QY 1 ATGGGCGCCGCGCCAGCATCTCTGGCGGCGAGAAAGCTGGGAGAAAGATCCGC 60
DB 1 ATGGGCGCCGCGCCAGCATCTCTGGCGGCGGCAAGCTGGGAGAAAGATCCGC 60
QY 61 CTGGCGCCCGCGGCGCAAGAAACACTACATGCTGAAGCACTGGTGTGGGCGAGCCGCGAG 120
DB 61 CTGGCGCCCGCGGCGGCGCAAGCACTACATGCTGAAGCACTGGTGTGGGCGAGCCGCGAG 120
QY 121 CTGGAGGGCTTGGCCCTGAACCCCGCCCTGCTGGAGACCGCCGAGGCTGCAAGCAGATC 180
DB 121 CTGGAGGGCTTGGCCCGGCGGCGCAAGCACTACATGCTGAAGCACTGGTGTGGGCGAGCCGCGAG 180
QY 181 ATGAAGCAGCTGCAGCCCGCCCTGCAGACCGGCGCACCGAGAGCTGCCAGCCTGACAC 240
DB 181 ATCAAGCAGCTGCAGCCCGCCCTGCAGACCGGCGCACCGAGAGATCCGCACTGTTCAAC 240
QY 241 ACCGTGGCCACCCCTGTACTGGCTGCACCGCGGCGCATCGAGTCCGCGACACCAAGGAGGCC 300
DB 241 ACCGTGGCCACCCCTGTACTGGCTGCACCAAGGCGCATCGAGTCCGCGGACACCAAGGAGGCC 300
QY 301 CTGGACAAGATCGAGGAGGAGCAGAAACAAGTCCCGAGCAGACACCCACAGGCCAAGGAG 360
DB 301 CTGGACAAGATCGAGGAGGAGCAGAAACAAGTCCCGAGCAGACACCCACAGGCCAAGGAG 360
QY 361 GCCGACGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGACCTGCAGGGCCAGATGGTG 420
DB 361 GCCGACAAAGAGGTGAGCCAGAACTACCCCATCGTGCAGACCTGCAGGGCCAGATGGTG 420
QY 421 CACCAAGGCCATCAGCCCCCGCACCCCTGAACGCTGGGTGAAGGTGATCGAGGAGAAGGCC 480
DB 421 CACCAAGGCCATCAGCCCCCGCACCCCTGAACGCTGGGTGAAGGTGATCGAGGAGAAGGCC 480
QY 481 TTCAGCCCCGAGGTGATCCCATGTTTCAACCGCCCTGAGCGAGGGCGCCACCCCGCAGGAC 540
DB 481 TTCAGCCCCGAGGTGATCCCATGTTTCAACCGCCCTGAGCGAGGGCGCCACCCCGCAGGAC 540
QY 541 CTGAACACCATGTTGAACACCGCTGGCGGCGCACCCAGGCCCATGCAGATGCTGAAGGAC 600
DB 541 CTGAACACCATGCTGAACACCGCTGGCGGCGCACCCAGGCCCATGCAGATGCTGAAGGAC 600
QY 601 ACCATCAACGAGGAGGCCGCGCAGTGGGACCGCCTGCACCGCGTGCAGGCGGCCCCCGTG 660
DB 601 ACCATCAACGAGGAGGCCGCGCAGTGGGACCGCCTGCACCGCGTGCAGGCGGCCCCCGTG 660
QY 661 GCCCGCGGCGCAGATGCGGACGCCCGCGGCGCAGCGACATCGCGGCGCCACGACGACCCCTG 720

||||||| 841 CTGGACATCAAGCAGGCGCCCAAGGAGCCCTTCGCGACTACGTGGACCGCTTCTTCAAG 900
QY 898 ACCCTGCGCGCGAGCAGGCGCCACCCAGGACCTGAAGAACTGGATGACCGAGACCTGCTG 957
Db 901 ACCCTGCGCGCGAGCAGGCGCCACCCAGGAGGTGAAGAACTGGATGACCGAGACCTGCTG 960
QY 958 GTGCAGAAAGCGCAACCCCGACTGCAAGACCATCTCGCGCTCTCGGCGCGCGCGCGCAC 1017
Db 961 GTGCAGAAAGCGCAACCCCGACTGCAAGACCATCTCGCGCTCTCGGCGCGCGCGCGCAC 1020
QY 1018 CTGGAGGAGATGATGACCGCTGCGCAGGGCGGTGGCGCGCGCGCGCGCGCGCGCGGTG 1077
Db 1021 CTGGAGGAGATGATGACCGCTGCGCAGGGCGGTGGCGCGCGCGCGCGCGCGCGCGGTG 1080
QY 1078 CTGGCGGAGGGGATGAGCCAGGCGCAACAGCGGTGAACATCATGATGCAGAGAGCAACTTC 1137
Db 1081 CTGGCGGAGGGGATGAGCCAGGCGCAACACCC---AGCGTATGATGCAGAGAGCAACTTC 1137
QY 1138 AAGGCGCCCGCGCGCAACGTCAAGTGTCTCAACTGCGGCAAGAGGGCGCCACATCGCCAA 1197
Db 1138 AAGGCGCCCGCGCGCATCTCAAGTGTCTCAACTGCGGCAAGAGGGCGCCACATCGCCCGC 1197
QY 1198 AACTGCGCGCGCGCGCGCAAGAGGGGTCTGTGAAGTGGGCAAGAGGGGTACACAGATG 1257
Db 1198 AACTGCGCGCGCGCGCGCAAGAGGGGTCTGTGAAGTGGGCAAGAGGGGTACACAGATG 1257
QY 1258 AAGGACTGCACCGAGCGCGCGCAAGTCTCTGGGCAAGATCTGGCGCAGGCGCAAGGGG 1317
Db 1258 AAGGACTGCACCGAGCGCGCGCAAGTCTCTGGGCAAGATCTGGCGCAGGCGCAAGGGG 1317
QY 1318 CGCGCGCGCAACTCTCTGCAGAGCGCGCAGCGGAGCGCGCGCGCGCGCGCGCGCGCG 1377
Db 1318 CGCGCGCGCAACTCTCTGCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1359
QY 1378 CGCGCGCGCGAGAGCTTCGCTTCGAGGAGACACCGCGCGCGCGCGCGCGCGCGCGCG 1437
Db 1360 CGCGCGCGCGAGAGCTTCGCTTCGAGGAGACACCGCGCGCGCGCGCGCGCGCGCGCG 1419
QY 1438 GACCGCGAGCCCTACCGCGGAGCGCGCGCTGACCGCGCTGCGCAGGCTGTTGCGCAGCGCGCG 1497
Db 1420 G-----ACCGCGAGACCTGACCGAGCGCTGAAGAGCGCTGTGGCGCAAGCGCG 1467
QY 1498 CTGAGCCAGTAA 1509
Db 1468 CTGAGCCAGTAA 1479

RESULT 8
AX468543
LOCUS AX468543 1479 bp DNA linear PAT 16-JUL-2002
DEFINITION Sequence 63 from Patent WO0226209.
ACCESSION AX468543
VERSION AX468543.1 GI:21901373
KEYWORDS
SOURCE Human immunodeficiency virus 1 (HIV-1)
ORGANISM Human immunodeficiency virus 1
Viruses: Retroviral viruses; Retroviridae; Lentivirus; Primate
lentivirus group.
REFERENCE 1
AUTHORS O'Hagan,D., Otten,G., Donnelly,J.J., Polo,J.M., Barnett,S.,
Singh,M., Ulmer,J. and Dubensky,J.W.
TITLE Microparticles for delivery of the heterologous nucleic acids
JOURNAL Patent: WO 0226209-A 63 04-APR-2002;
CHIRON CORPORATION (US)
FEATURES
source Location/Qualifiers
1..1479
/organism="Human immunodeficiency virus 1"
/mol_type="genomic DNA"
/db_xref="taxon:11676"
BASE COUNT 325 a 529 c 463 g 162 t
ORIGIN

Query Match: 84.6%; Score 1276.8; DB 6; Length 1479;
Best Local Similarity 92.5%; Pred. No. 2e-125;
Matches 1399; Conservative 0; Mismatches 77; Indels 36; Gaps 4;
QY 1 ATGGGCG 60
Db 1 ATGGGCG 60
QY 61 CTGGCG 120
Db 61 CTGGCG 120
QY 121 CTGGAGGCGCTTCGCGCTGAACCCCGCGCTGCTGAGAGACCGCGCGCGCGCGCGCGCG 180
Db 121 CTGGAGGCGCTTCGCGCTGAACCCCGCGCTGCTGAGAGACCGCGCGCGCGCGCGCGCG 180
QY 181 ATGAGCAGCIGCAGCG 240
Db 181 ATCCGCGAGCTGCACCG 240
QY 241 ACCGTGCG 300
Db 241 ACCGTGCG 300
QY 301 CTGGAACAAGATCGAGGAGGAGCAGAAACAAAGTCCGAGCAGAAAGACCGCGCGCGCGCG 360
Db 301 CTGGAACAAGATCGAGGAGGAGCAGAAACAAAGTCCGAGCAGAAAGATCCGAGCGCGCGCG 360
QY 361 GCGGAC--GGCAAGGTGAGCG 417
Db 361 GCGGACAAGGGCAAGGTGAGCG 420
QY 418 GTGACACAGGCG 477
Db 421 GTGACACAGGCG 480
QY 478 GCCTTCAGCG 537
Db 481 GCCTTCAGCG 540
QY 538 GACCTGAACACAGATGTGAACACCGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 597
Db 541 GACCTGAACACAGATGTGAACACCGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
QY 598 GACACCATCAACGAGGAGGCG 657
Db 601 GACACCATCAACGAGGAGGCG 660
QY 658 GTGGCG 717
Db 661 ATCG 720
QY 718 CTGAGGAGCAGATCGCTGGATGACCGAGCAACCGCGCGCGCGCGCGCGCGCGCGCGCG 777
Db 721 CTGAGGAGCAGATCGCTGGATGACCGAGCAACCGCGCGCGCGCGCGCGCGCGCGCGCG 780
QY 778 AAGCGGTGGATCATCTGGCGCTGAACAAGATCGTGGGATGTACAGCGCGCGCGCGCGCG 837
Db 781 AAGCGGTGGATCATCTGGCGCTGAACAAGATCGTGGGATGTACAGCGCGCGCGCGCGCG 840
QY 838 CTGGACATCG 897
Db 841 CTGGACATCAAGCAGGCG 900
QY 898 ACCCTGCG 957
Db 901 ACCCTGCG 960
QY 958 GTGCAGAAAGCGCAACCCCGACTGCAAGACCATCTCTCGCGCTCTCGGCGCGCGCGCG 1017
Db 961 GTGCAGAAAGCGCAACCCCGACTGCAAGACCATCTCTCGCGCTCTCGGCGCGCGCGCG 1020
QY 1018 CTGGAGGAGATGATGACCGCTGCGCAGGGCGGTGGCGCGCGCGCGCGCGCGCGCGGTG 1077


```
QY 1378 CCCCCCGGAGAGCTCCGCTTCAGGAGACACACCCCGCCGCGCCCAAGCAGGACGACG 1437
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1360 CCCCCCGGAGAGCTCCGCTTCAGGAGACACACCCCGCCGCGCCCAAGCAGGACGACG 1439
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 1438 GACCGGAGCCCTACCGGAGCCCTGACCGCCCTGCGCAGCCTGTCGCGACGCGCC 1497
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1420 G-----ACCGGAGACCCCTGACCGCCCTGAGAGCCTGCTTCGCGACGACCC 1467
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 1498 CTGAGCCAGTAA 1509
    '|||||:|||||:
Db 1468 CTGAGCCAGTAA 1479
    '|||||:|||||:

RESULT 11
LOCUS AX149648 4288 bp DNA linear PAT 08-JUN-2001
DEFINITION Sequence 2 from Patent WO0136614.
ACCESSION AX149648
VERSION AX149648.1 GI:14348047
KEYWORDS Human immunodeficiency virus
SOURCE Human immunodeficiency virus
ORGANISM Human immunodeficiency virus
Virus; Retroid viruses; Retroviridae; Lentivirus; Primate
lentivirus group.
REFERENCE
1 Shao,Y., Wagner,R., Wolf,H. and Graf,M.
The genome of the hiv-1 inter-subtype (c/b') and use thereof
Patent: WO 0136614-A 2 25-MAY-2001;
Geneart GMRH Gesellschaft fuer angewandte Biotechnologie (DE);
Shao, Yiming (CN)
FEATURES
    source
        1..4288
            /organism="Human immunodeficiency virus"
            /mol_type="genomic DNA"
            /db_xref="taxon:12721"
BASE COUNT 1026 a 1358 c 1397 g 507 t
ORIGIN
```

```
Query Match 79.7%; Score 1202.2; DB 6; Length 4288;
Best Local Similarity 88.5%; Pred. No. 1e-117;
Matches 1335; Conservative 0; Mismatches 153; Indels 21; Gaps 2;
```

```
QY 1 ATGGGCGCGCGCGCCACATCCTCGCGGGGAGAGCTGGACAACTGGGAGAGATCCGC 60
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 13 ATGGGCGCGCGCGCCACATCCTCGCGGGGAGAGCTGGACAACTGGGAGAGATCCGC 72
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 61 CTGCGCGCGCGCGCGAGAGAGCACTACATCCTGAACAGCTGTGTGGCGAGCGGAG 120
    '|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 73 CTGAGGCGCGCGCGCGAGAGCACTACATCCTGAACAGCTGTGTGGCGAGCGGAG 132
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 121 CTGAGGCGCTTCGCCCTGAACCCCGGCTGCTGGAGACCGCGAGGAGCTGCAACAG 240
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 193 ATGAGGAGCTGCAGAGCGCCCTGCGACACCGGACCGGAGGAGCTGAGGAGCTGTTCAAC 252
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 241 ACCGTGCGCACCCCTGTACTGCGTGCAGCGCGCGCATCGAGGTCCCGGACACCGAGGCG 300
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 253 ACCGTGCGCACCCCTGTACTGCGTGCAGACCGGAGCTGACCTGAGGACACCGAGGAGGCC 312
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 301 CTGGACAGATCGAGGAGGAGCAGACAAAGTCCACAGAGAGACCCAGCAGGCGCAAGGAG 360
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 313 CTGGACAGATCGAGGAGGAGCAGACAAAGATCCACAGAGAGACCCAGCAGGCGCAAGGAG 372
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 361 GCCGACGGCAAGGTGAGCCAGAACTACCCCATCGTCAGAACCTGCGAGGCGCCAGATGGTG 420
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 373 GCCGACGGCAAGGTGAGCCAGAACTACCCCATCGTCAGAACCTGCGAGGAGCTGCGAGGCGCAGATGGTG 432
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 421 CACGAGGCGATCAGCCCGCGCACCCCTGAACGCGCTGGGTGAAGGTGATCGAGGAGAGGCC 480
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 433 CACGAGGCGATCAGCCCGCGCACCCCTGAANTGCAIGGGTGAAGGTGGTGGAGGAGAGGCC 492
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

```
QY 481 TTACGCGCGCGAGGTGATCCCATCTTCACCGCGCTGAGCGAGCGCGCCACCCCGCAGGAC 540
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 493 TTACGCGCGCGAGGTGATCCCATCTTCAGCGCGCTGAGCGAGCGCGCCACCCCGCAGGAC 552
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 541 CTGAACACGATGTTGAACACCGTGGGCGCCACAGCGCGCATGACGATGCTCAAGGAC 600
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 553 CTGAACACCATGCTGAACACCGTGGGCGCCACAGCGCGCATGACGATGCTCAAGGAC 612
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 601 ACCATCAACGAGGAGCGCGCGAGTGGGACCGCTTCACCGCGTGCAGGCGCGCCCGTG 660
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 613 ACCATCAACGAGGAGCGCGCGAGTGGGACAGCTGCACCCCGTGCACCGCGCGCCCATC 672
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 661 GCCCGCGCGCAGATGCGCGACCCCGCGCGGACGACATCGCGCGCGCCACAGCACCCCTG 720
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 673 GCCCGCGCGCAGATGAGGGAGCCCGAGGGGACGACATCGCGCGGACACAGCAACCTG 732
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 721 CAGGAGCAGATCGCCTGGATGACGACGACACCCCGCGCTGCGCGTGGCGGACATCTACAG 780
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 733 CAGGAGCAGATCGCCTGATGACGACGACACCCCGCGCTGCGCGTGGCGGACATCTACAG 792
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 781 CGGTGATCATCTCTGGCGCTGAACAAGATCGTGGGATGTACAGCCCGCTGAGCATCCTG 840
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 793 AGGTGATCATCTCTGGGTTTAACAAGATCGTGGGATGTACAGCCCGCAGCATCCTG 852
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 841 GACATCGCGCAGGCGCGCGCGCGCGCTTCGCGGACTACGTGGACCGCTTCTTCAAGACC 900
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 853 GACATCAAGCAGGCGCGCGCGCGCGCTTCAGGAGCTACGTGACAGGTTCTTCAAGACC 912
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 901 CTGCGCGCGGAGCAGGCGCGCGCGCGCGCTGAAGAAGTGGATGACCGGACCGCTGCTGGTG 960
    '|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 913 CTGAGGCGGAGCAGGCGCGCGCGCGCGCTGAAGAAGTGGATGACCGGACCGCTGCTGGTG 972
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 961 CAGAACGCCAACCCCGACTGCAAGACCATCTCGCGCTCTCGCGCGCGCGCGCGCACCCCTG 1020
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 973 CAGAACGCCAACCCCGACTGCAAGACCATCTCGAGGCGCTTGGCGCGCGCGCGCGCAGCAIC 1032
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 1021 GAGGAGATGATGACCGCTGCCAGGCGGTGGCGCGCGCGCGCGCACAGCGCGCTGCTG 1080
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1033 GAGGAGATGATGACCGCTGCCAGGCGGTGGCGCGCGCGCGCGCACAGCGCGCTGCTG 1092
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 1081 GCCGAGGCGATGAGCGAGCGCGCGCGCGCGCTGAACATCATGTATGAGAGAGCAACTTCAAG 1140
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1093 GCCGAGGCGATGAGCGAGCGCGCGCGCGCGCTGAACATCATGTATGAGAGAGCAACTTCAAG 1149
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 1141 GCGCGCGCGCGCGCGCGCTCAAGTCTCAACTCGCGCGAGGAGCGCGCATCGCGCGCAAGAAC 1200
    '|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1150 GCGAGCAAGAGGATGCTGAAGTCTTCAACTCGCGCGAGGAGCGCGCATCGCGCGCAAGAAC 1209
    '|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 1201 TGCGCGCGCGCGCGCGCAAGAGGGCTGCTGGAAGTGGCGAGAGGCGCGCATCGAGATGAAG 1260
    '|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1210 TGCGCGCGCGCGCGCGCAAGAGGGCTGCTGGAAGTGGCGAGAGGCGCGCATCGAGATGAAG 1269
    '|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 1261 GACTGCGCGCGCGCGCGCGCGCTTCCCTGGGCAAGATCTGGCGCGCGCGCGCGCGCGCGCG 1320
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1270 GACTGCGCGCGCGCGCGCGCGCTTCCCTGGGCAAGATCTGGCGCGCGCGCGCGCGCGCGCG 1329
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 1321 CCGCGCAACTTCTGCGAGAACCCCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1330 CCGCGCAACTTCTGCGAGAACAG-----GCGCGAGCGCGCGCGCGCGCGCGCGCG 1371
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 1381 CCGCGCGGAGGCTTCCGCTTCGAGGAGACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1440
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1372 CCGCGGAGAGGCTTCAGGTTTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1431
    '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 1441 CGCGAGCGCTTACCGCGGAGCGCGCTGACCGCGCGCTGCGGAGCGCTTTCGGCAGCGCGCGCG 1500
    '|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1432 ATCGACAGGAGGCTGTACCCCGCTGACCGAGCGCTGAGAGGCGCTTTCGGCAGCGCGCGCG 1491
    '|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 1501 AGCCAGTAA 1509
    '|||||:|||||:
Db 1492 AGCCAGGAA 1500
    '|||||:|||||:
```



```

RESULT 12
AF201927
LOCUS      Synthetic construct gag protein gene, complete cds.          linear      SYN 16-MAR-2000
ACCESSION  AF201927
VERSION    AF201927.1  GI:7248702
KEYWORDS
SOURCE     synthetic construct
ORGANISM   synthetic construct
            artificial sequences.
REFERENCE  1 (bases 1 to 1509)
            zur Megede,J., Chen,M.C., Doe,B., Schaefer,M., Greer,C.E.,
            Selby,M., Otten,G.R. and Barnett,S.W.
            Increased expression and immunogenicity of sequence-modified human
            immunodeficiency virus type 1 gag gene
            J. Virol. 74 (6). 2628-2635 (2000)
JOURNAL    20148954
MEDLINE    10684277
PUBMED
REFERENCE  2 (bases 1 to 1509)
            zur Megede,J. and Barnett,S.W.
            Direct Submission
            Submitted (04-NOV-1993) Vaccines, Chiron Corporation, 4560 Horton,
            Emeryville, CA 94608, USA
FEATURES   Location/Qualifiers
            1..1509
                /organism="synthetic construct"
                /mol_type="genomic DNA"
                /db_xref="taxon:32630"
            1..1509
                /note="HIV-1SF2 p55gag"
                /codon_start=1
                /translation="MGARASVLSGGELDKWEKIRLPGGKKKKYKLLKHIVWASRELERF
                AVNPLLETSECRQILGQLPSLQTGTSEELPSLYNTVATLYCVHQRIDVADTKEALE
                KJEEONKSKKAAQAAAAGTNSOVSPYIVONLQGMVHOAISPRTINANVKV
                VEEKAFSEVTPMFSAISEGATPDLLNTMLNTVGGHQAAMQMLKETINEEAFWDRVH
                PVHAGPIAGQMREPRGSDIAGTTSILOEQIGWMTNPPIPVGEIYKRWIILGINKLY
                RMVSPTSLIDIRQPKPEFRYVDRFYKTLRAEQASQDVKNWMTETLLIVONANPDCKT
                ILKALGPAATLEEMMTACQGVGGPGHKARVLAEMSQVNPATIMMORGFNMQRTY
                KCFNCGKEGHTARNCRAPRKKGCRWCRGREGHMKDCTEROANFLGKIWPSTYKGRGNGF
                LOSRPEPTAPDEESFRGEEKTTPSQKQEPIDKELYPLISLRSFGNDPSSQ"
            328 a      543 c      479 g      159t
BASE COUNT 328 a      543 c      479 g      159t
ORIGIN
Query Match      79.6%; Score 1201; DB 12; Length 1509;
Best Local Similarity 86.5%; Pred. No. 1,8e-117;
Matches 1351; Conservative 0; Mismatches 140; Indels 36; Gaps 3;
QY      1  ATGGGCGCGCGCGGCAGCATCTGCGCGCGGAGAGAGCTGGACAAAGTGGGAGAGATCGCG 60
Db      1  ATGGGCGCGCGCGCGGCAGCTGCTGAGCGCGGCGGAGTGGACAAAGTGGGAGAGATCGCG 60
QY      61  CTGCGCGCGCGCGGCAAGAAGCACTACATGCTGAAGCACTTGTTGGGCCAGCGCGAG 120
Db      61  CTGCGCGCGCGCGGCAAGAAGAGTACAAGCTGAAGCACTGTTGGGCCAGCGCGAG 120
QY      121  CTGGAGGCGCTTCGCCCTGAACCCCGGCTGCTGGAGACCGCGGAGGCTGCAAGCAGATC 180
Db      121  CTGGAGGCGCTTCGCCGTGAACCCCGGCTGCTGGAGACCGAGGCTGCGCCAGATC 180
QY      181  ATGAAGCAGCTGCAGCCCGCCCTGCAGACCGGCGACCGAGGAGCTGCGCAGCCTGTACAAC 240
Db      181  CTGGCCAGCTGCAGCCCGAGCCTGCAGACCGGCGAGCGAGGAGCTGCGCAGCCTGTACAAC 240
QY      241  ACCGTGGCACCCCTGTACTCGTGTCACCCCGGCGATCGAGGTCCGGGACACCAAGGAGGCC 300
Db      241  ACCGTGGCACCCCTGTACTCGTGTCACCGAGCGGATCGAGCTCAAGGACACCAAGGAGGCC 300
QY      301  CTGGACAAATCGAGGAGGAGCAGACAAAGTCCAGCAGAGAGACCCAGCAGGCCAAGGAG 360

```


Db 1363 CCGAGCCACCCGCCCCCGGAGGAGGCTTCGGGAGGAGAGACCAACCCCG 1422

QY 1423 AAGCAGAGCCCAAGAGCGGAGCCCTACCGGAGGCCCCGTGAGCGCCCTGCGAGCTTG 1482

Db 1423 AGCCAGAGCAGGAGCCCATCCAGAGGAGCTGTACCCCTGACAGCCTGCGAGGCTG 1482

QY 1483 TTCGGCAGCGGCCCTCTGAGCCAGTAA 1509

Db 1483 TTCGCAAGCAGCCCGCAGCAGCCAGTAA 1509

RESULT 13

AX427930

LOCUS AX427930 9166 bp DNA linear PAT 20-JUN-2002

DEFINITION Sequence 168 from Patent WO0232943.

ACCESSION AX427930

VERSION AX427930.1 GI:21538017

KEYWORDS

SOURCE synthetic construct

ORGANISM synthetic construct

artificial sequences.

REFERENCE 1

AUTHORS Huang,Y. and Nabel,G.J.

TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for genetic immunization

JOURNAL Patent: WO 0232943-A 168 25-APR-2002;

GOVERNMENT OF THE UNITED STATES (US)

FEATURES

Source Location/Qualifiers

1..9166

/organism="synthetic construct"

/mol_type="genomic DNA"

/db_xref="taxon:32630"

/note="plasmid pVR1012x/s containing HIV genes"

BASE COUNT 2227 a 2662 c 2513 g 1764 t

ORIGIN

Query Match 78.6%; Score 1186; DB 6; length 9166;

Best Local Similarity 87.7%; Pred. No. 4.1e-115;

Matches 1326; Conservative 0; Mismatches 165; Indels 21; Gaps 2;

QY 1 ATGGGCGCCCGCGCCAGCATCTCTGCGCGGCGAGAGACCTGGGAGAGTGGGAGAGATCCGC 60

Db 1880 ATGGGCGCCCGCGCCAGCATCTCTGCGCGGCGGAGAGTGGGAGAGTGGGAGAGATCAAG 1939

QY 61 CTGCGCGCCCGCGCGGAGAGCAGCAGTACATCTGAGGAGCAGCTGCTGAGGAGCAGCTGCTGAGGAGCAGCAGCAG 120

Db 1940 CTGAAGCGCGCGCGGAGAGAGCAGCAGTACATGATGAGGAGCAGCTGCTGAGGAGCAGCAGCAG 1999

QY 121 CTGAGGAGCTTGGCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 180

Db 2000 CTGAGGAGCTTGGCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2059

QY 181 ATGAAGCAGCTGCAGCGCCCTCTGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 240

Db 2060 ATGAAGCAGCTGCAGCGCCCTCTGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2119

QY 241 ACCGTGCCACCTGTACTGCGTGCAGCGCCCTGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 300

Db 2120 ACCGTGCCACCTGTACTGCGTGCAGCGCCCTGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2179

QY 301 CTGGACAAGATCGAGGAGGAGCAGAACAAAGTCCAGCAGAGAGAGAGAGAGAGAGAGAGAGAG 360

Db 2180 CTGGACAAGATCAAGGAGGAGCAGAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2239

QY 361 GCCGAC---GGCAAGTGAGCCAGAACTACCCCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 417

Db 2240 GCCGACAAGGCAAGTGAGCCAGAACTACCCCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2299

QY 418 GTGCACAGGCCCATCAGCGCCCGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGAG 477

Db 2300 GTGCACAGGCCCATCAGCGCCCGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGAG 2359

QY 478 GCCTTCAGCCCCGAGGTGATCCCCATGTTACCCGCGCTGAGCGAGGCGCGCCACCCCGCAG 537

Db 2360 GCCTTCAGCCCCGAGGTGATCCCATGTTACAGCCCTGAGGAGGCGCCACCCCGCAG 2419

QY 538 GACCTGAACACAGATGTTGAACACCGTGGCGGCGCCACAGGCGCCATGCGAGATGCTGAAG 597

Db 2420 GACCTGAACACAGATGTTGAACACCGTGGCGGCGCCACAGGCGCCATGCGAGATGCTGAAG 2479

QY 598 GACACCATCAACGAGGAGCGCGGAGTGGAGCCGCCCTGACCCCGTGCAGGCGCGCCCG 657

Db 2480 GACACCATCAACGAGGAGCGCGGAGTGGAGCCGCCCTGACCCCGTGCAGGCGCGCCCG 2539

QY 658 GTGGCCCCCGGCGCAGATGCGGCGACCCCGCGGCGAGCGACATGCGCGGCGCCACAGCACC 717

Db 2540 GTGGCCCCCGGCGCAGATGCGGCGAGCCCGCGGCGAGCGACATGCGCGGCGCCACAGCACC 2599

QY 718 CTGCAGGAGCAGATGCGCTGGATGACCGAGCAGACCCCGCGCGCGTGGCGGACATCTAC 777

Db 2600 CTGCAGGAGCAGATGACCTGGATGACCAACACACCCCGCGCGCTGGCGGACATCTAC 2659

QY 778 AAGCGGTGGATCATCTCTGGCCTGAACAAGATGCTGCGGATGTACAGCCCGCTGAGCATC 837

Db 2660 AAGCGCTGGATCATCTCTGGCCTGAACAAGATGCTGCGGATGTACAGCCCGCTGAGCATC 2719

QY 838 CIGACATCCGCGAGGCGCCCAAGGAGCCCTTCCGCGACTAGCTGGAGCCGCTTCTTCAAG 897

Db 2720 CTGGACATCAAGCAGGCGCCCAAGGAGCCCTTCCGCGACTAGCTGGAGCCGCTTCTTCAAG 2779

QY 898 ACCCTGCGCGCGGAGCAGGCGCCACCGAGGAGCTGAAGAAGTGAATGACCGAGACCCCTGCTG 957

Db 2780 ACCCTGCGCGCGGAGCAGGCGCCACCGAGGAGCTGAAGAAGTGAATGACCGAGACCCCTGCTG 2839

QY 958 GTGCAGAACGCCAACCCCGACTCAAGAGCCATCTCTGCGCGCTCTCGGCGCGCGCGCCACC 1017

Db 2840 GTGCAGAACGCCAACCCCGACTCAAGAGCCATCTCTGCGCGCTCTCGGCGCGCGCGCCACC 2899

QY 1018 CTGGAGGAGATGATGACCGCTTCCAGGCGCTGCGCGCGCGCCACAGCGCCCGCGTG 1077

Db 2900 CTGGAGGAGATGATGACCGCTTCCAGGCGCTGCGCGCGCGCCACAGCGCCCGCGTG 2959

QY 1078 CTGGCGGAGGCGATGAGCCAGGCGCAACAGCGTGAACATCATGATGAGAGAGCAACTTC 1137

Db 2960 CTGGCGGAGGCGATGAGCAAGGTGAACAACACCAACATCATGATGAGCGAGCAACTGC 3019

QY 1138 AAGGCGCCCGCGGCGCAACGTCAATGCTTCAACTGCGGCAAGGAGGCGCCACAG 1197

Db 3020 AAGGCGCCCGCGGCGCAACGTCAATGCTTCAACTGCGGCAAGGAGGCGCCACAG 3079

QY 1198 AACTGCGCGCGCGCGCGCAAGAGGCGTGTGGAAGTGTGCGCAAGGAGGCGCCACAGATG 1257

Db 3080 AACTGCGCGCGCGCGCGCAAGAGGCGTGTGGAAGTGTGCGCAAGGAGGCGCCACAGATG 3139

QY 1258 AAGGACTGCACCGAGCGCGCGCAACTTCTTGGGCAAGATGCTGCGCCAGCGCAAGGGC 1317

Db 3140 AAGGACTGCACCGAGCGCGCGCAACTTCTTGGGCAAGATGCTGCGCCAGCGCAAGGGG 3199

QY 1318 CGCCCGCGCAACTTCTTGCAGACCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1377

Db 3200 AGGCCAGGGAATTTCTTCAGAGCAG-----ACGAGAGCGCAACAGCC 3241

QY 1378 CCGCCCGCGCGAGAGCTTCCGCTTCGAGGAGAGACACCCCGCGCGCGCGCGCGCGCGCGCG 1437

Db 3242 CCACGAGAAGAGAGCTTCAGGTTTGGGGAAGAGACACCACTCCCTCTCAGAGAGCAGGAG 3301

QY 1438 GACCGCGAGCCCTACCGCGAGCGCCCTGACCGCGCTTGGAGCGCTGTTGGCAGCGCGCC 1497

Db 3302 CCGATAGACAAGAACTGTATCCTTTAGCTTCCCTCAGATCACTCTTTGGCAGCGACCCCG 3361

QY 1498 CTGAGCCAGTAA 1509

Db 3362 TCCTCACAATAA 3373

AX427936
LOCUS AX427936 9788 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 174 from Patent WO0232943.
ACCESSION AX427936
VERSION AX427936.1 GI:21538023
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE
1 Huang, Y. and Nabel, G. J.
AUTHORS Modifications of hiv env, gag, and pol enhance immunogenicity for
TITLE genetic immunization
JOURNAL Patent: WO 0232943-A 174 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
FEATURES
Location/Qualifiers
source 1..9788
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="plasmid pVRI012x/s containing HIV genes"
BASE COUNT 2377 a 2817 c 2695 g 1894 t
ORIGIN
Query Match 77.6%; Score 1170.8; DH 6; Length 9788;
Best Local Similarity 87.6%; Pred. No. 1.6e-114;
Matches 1324; Conservative 0; Mismatches 162; Indels 26; Gaps 3;
QY 1 ATGGGGCGCGCGCCAGCATCTCTGCGCGGCGAGAGCTGGACAGTGGGAGAGATCCGC 60
Db |||||
1880 ATGGGGCGCGCGCCAGCATCTCTGCGCGGCGGCGAGCTGGAGCTGGGAGCGCATCAAG 1939
QY 61 CTGCGCCCGCGGCGGCAAGACACTACATGCTGAAGCACCTGGTGGTGGCGCGCGAG 120
Db |||||
1940 CTGAAGCGCGGCGGCAAGACACTACATGATGAAGCACCTGGTGGTGGCGCGAG 1999
QY 121 CTGAGGGCTTCGCCCTGAACCCCGGCTGCTGGAGACCGCGGAGGCTGCAAGCAGATC 180
Db |||||
2000 CTGAGGCTTCGCCCTGGACCGCGGCTGCTGGAGACCGCGGAGGCTGCAAGCAGATC 2059
QY 181 ATGAAGAGCTGCAGCGCGCGCTGCAAGACCGCGGAGGAGCTGGCAGCGCTGTACAAC 240
Db |||||
2060 ATGAAGAGCTGCAGCGCGCGCTGCAAGACCGCGGAGGAGCTGGATCAGCTGCACAAAC 2119
QY 241 ACCGTGGCACCCCTGTACTGCTGCACCGCGGCTGCTGAGGCTCGGAGACCAACAGGCG 302
Db |||||
2120 ACCGTGGCACCCCTGTACTGCTGCAGAGAACTGACGCTGCGGACACCAAGAGGCG 2179
QY 301 CTGGACAGATCGAGGAGGAGCAAGAACTCCAGAGAGACCGAGGAGCGCAAGAG 360
Db |||||
2180 CTGGACAGATCAAGGAGGAGGAGCAAGAAAGGCGGAGAGAGACCGAGGAGGCG 2239
QY 361 GCGGAC---GGCAAGTGAGCGGAGAACTACCCCATGCTGCAGAACTGCAGGCGGAGATG 417
Db |||||
2240 GCGGACAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2299
QY 418 GTGCACAGGCGGATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 477
Db |||||
2300 GTGCACAGGCGGATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2359
QY 478 GCCTTCAGCGCGGAGGTGATCCCGCATGTTACCGCGCTGAGCGGCGCGCGCGCGCGCG 537
Db |||||
2360 GCCTTCAGCGCGGAGGTGATCCCGCATGTTACCGCGCTGAGCGGCGCGCGCGCGCGCG 2419
QY 538 GACCTGACACGATGTTGAACACCGCTGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 597
Db |||||
2420 GACCTGACACGATGCTGAACACCGCTGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 2479
QY 598 GACACCATCAACGAGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 657
Db |||||
2480 GACACCATCAACGAGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2539
QY 658 GTGGCG 717

Db 2540 GTGGCG 2599
QY 718 CTGCAGGAGCAGATCGCCCTGGATGATGATGATGATGATGATGATGATGATGATGATGAT 777
Db |||||
2600 CTGCAGGAGCAGATCACCTGGATGATGATGATGATGATGATGATGATGATGATGATGAT 2659
QY 778 AAGCGGTGATCATCTCTGGGCTGAACAAAGATCTGCGGATGATGATGATGATGATGATGAT 837
Db |||||
2660 AAGCGGTGATCATCTCTGGGCTGAACAAAGATCTGCGGATGATGATGATGATGATGATGAT 2719
QY 838 CTGGACATCCCG 897
Db |||||
2720 CTGGACATCAAGCAGGCG 2779
QY 898 ACCCTGCG 957
Db |||||
2780 ACCCTGCG 2839
QY 958 GTGCAGAACCGCAACCG 1017
Db |||||
2840 GTGCAGAACCGCAACCG 2899
QY 1018 CTGGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1077
Db |||||
2900 CTGGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2959
QY 1078 CTGGCGGAGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1137
Db |||||
2960 CTGGCGGAGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3019
QY 1138 AAGGGCG 1197
Db |||||
3020 AAGGGCG 3079
QY 1198 AACTGCG 1257
Db |||||
3080 AACTGCG 3139
QY 1258 AAGGACTGACCGGCG 1317
Db |||||
3140 AAGGACTGACCGGCG 3194
QY 1318 CGCCCG 1377
Db |||||
3195 AGGCCG 3236
QY 1378 CCG 1437
Db |||||
3237 CCACGAGAGAGAGCTTCAGGTTTGGGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3296
QY 1438 GACCGGAGCG 1497
Db |||||
3297 CCGATAGACAG 3356
QY 1498 CTGAGCGCGTAA 1509
Db 3357 TCGTCAACAATAA 3368
RESULT 15
AX306431
LOCUS AX306431 1539 bp DNA linear PAT 11-DEC-2001
DEFINITION Sequence 5 from Patent WO0188141.
ACCESSION AX306431
VERSION AX306431.1 GI:17645655
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1
AUTHORS Wagner, R., Graf, M., Deml, L. and Bieler, K.
TITLE Synthetic gapol genes and their uses

JOURNAL Patent: WO 0188141-A 5 22-NOV-2001

Genbank (DE)
Location/Qualifiers
1. 1539

FEATURES
Source

/organism="synthetic construct"

/mol_type="genomic DNA"

/db_xref="taxon:32630"

/note="sequence with optimized codons"

BASE COUNT 366 a 508 c 502 g 163 t

ORIGIN

Query Match 75.9%; Score 1145; DB 6; Length 1539;

Best Local Similarity 85.7%; Pred. No. 1.4e-111;

Matches 1319; Conservative 0; Mismatches 190; Indels 30; Gaps 5;

```
QY 1 ATGGGCGCGCGCCAGCATCTGCGCGGGGAGAGAGCTGGACACTGGGAGAGATCCGG 60
Db 1 ATGGGCGCGCGCCAGCGCTGCTGAGCGCGCGGAGCTGGACAGTATGAGCTGAG 60
QY 61 CTGGCGCGCGCGCGCGCAAGACACTACATGCTGAACCACTTGTGTGGGCGAGCGGAG 120
Db 61 CTGAGCGCGCGCGCGCGCAAGAGAGTATAGCTGAAGCACTATGCTGTGGCGCGAGGAG 120
QY 121 CTGGAGGCGCTGCGCCCTGAACCCCGCGCTGCTGGAGACCGCGCGAGGCTGCAGCAGATC 180
Db 121 CTGGAGAGGCTGCGCGCTGAACCCCGCGCTGCTGGAGACCGCGCGAGGCTGCAGCAGATC 180
QY 181 ATGAAGCAGCTGCAGCGCGCGCTGCGAGACCGCGCGAGGAGCTGCGCAGCCGTATCAAC 240
Db 181 CTGGCGCAGCTGCAGCGCGCGCTGCGAGACCGCGCGAGGAGCTGAGGAGCCTGTACAC 240
QY 241 ACCGTGGCGACCCCTGTACTGCGTGCAGCGCGCGCTGCGAGGCTGCGCGACACCAAGGCG 300
Db 241 ACCGTGGCGACCCCTGTACTGCGTGCAGCGCGCGCTGCGAGGCTGAGGAGCCTGTACAC 300
QY 301 CTGGACAGATCGAGGAGGAGCAGACAAAGTCCACAGAGAGAGCCAGCGAGGAGGAG 360
Db 301 CTGGACAGATCGAGGAGGAGCAGACAAAGTCCACAGAGAGAGCGCCAGCGAGGCGCG 360
QY 361 GCCGACG-----GCCAAGGTGAGCCAGACACACCCCATGTCAGAACTGCGAGGCG 411
Db 361 GACACCGCGCACAGCAGCGAGGTGAGCCAGAACTACCCCATGTCGACAGAACTCCAGGCG 420
QY 412 CAGATGTCACACGAGGAGCAGCGCGCGCGCATGTCACCCCATGTCAGAACTGATCGAG 471
Db 421 CAGATGTCACACGAGGAGCAGCGCGCGCGCATGTCACCCCATGTCAGAACTGATCGAG 480
QY 472 GAGAAGGCTTCAGCGCGCGCGCGCGCGCATGTCACCCCATGTCAGAACTGATCGAG 531
Db 481 GAGAAGGCTTCAGCGCGCGCGCGCGCGCATGTCACCCCATGTCAGAACTGATCGAG 540
QY 532 CCCCAGGACCTGAACAGGATGTTGAACACCGTGGCGCGCGCGCGCGCGCGCGCGCGCG 591
Db 541 CCCCAGGACCTGAACAGGATGTTGAACACCGTGGCGCGCGCGCGCGCGCGCGCGCGCG 600
QY 592 CTGAAGGACAGCATCAACAGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 651
Db 601 CTGAAGGAGACCATCAACAGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
QY 652 GGCGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 711
Db 661 GGCGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
QY 712 AGCACCCCTGCAGGAGGAGATCGCCCTGGATGACCAAGCAACCCCGCGTGGCGCGGAC 771
Db 721 AGCACCCCTGCAGGAGGAGATCGCCCTGGATGACCAAGCAACCCCGCGTGGCGCGGAA 780
QY 772 ATCTACAGCGGTGGATCATCTCTGGGCGCTGAACAGATCGTGGCGATGTACAGCCCGCTG 831
Db 781 ATCTACAGAGGTGGATCATCTCTGGGCGCTGAACAGATCGTGGCGATGTACAGCCCGCTG 840
QY 832 AGCATCTCTGGACATCCCGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 891
```

Search completed: September 22, 2003, 20:56:49
Job time : 5718.74 secs

```
Db 841 AGCATCTCTGGATATCAGGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 900
QY 892 TTAACAGACCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 951
Db 901 TACAGACCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 960
QY 952 CTGCTGGTGCAGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1011
Db 961 CTGCTGGTGCAGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020
QY 1012 GCCACCTTGGAGGAGATGATGACCGCGCTGCCAGCGCGCGCGCGCGCGCGCGCGCG 1071
Db 1021 GCCACCTTGGAGGAGATGATGACCGCGCTGCCAGCGCGCGCGCGCGCGCGCGCGCG 1080
QY 1072 CGCGTGTGGCGCGAGGCGATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1128
Db 1081 AGGTGTGTGGCGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1140
QY 1129 AGCAACTTCAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1188
Db 1141 GGCAACTTCAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1200
QY 1189 ATGCGCAAGAACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1248
Db 1201 ACCGCGAGAACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1260
QY 1249 CACAGATGAAGGAGTGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1308
Db 1261 CACAGATGAAGGAGTGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1320
QY 1309 CACAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1362
Db 1321 TACAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
QY 1363 -----ACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1410
Db 1381 CTGAGAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1440
QY 1411 ACCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1470
Db 1441 ACCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1500
QY 1471 CTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1509
Db 1501 CTGAGGAGCGCTGTTCGGCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1539
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2003, 22:50:29 : Search time 358.76 Seconds
(without alignments)
9866.087 Million cell updates/sec

Title: US-09-475-704A-3
Perfect score: 1479
Sequence: 1 atgggccccgcgagcagcat.....acgacccccctgagccagtau 1479

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1660708 seqs, 122995015 residues

Total number of hits satisfying chosen parameters: 3321416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PCHCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1479	100.0	1479	12	US-09-967-464-63
2	1477.4	99.9	3152	12	US-10-190-435-18
3	1477.4	99.9	3462	12	US-10-190-435-16
4	1477.4	99.9	4419	12	US-10-190-435-19
5	1477.4	99.9	4483	12	US-10-190-435-35
6	1477.4	99.9	4606	12	US-10-190-435-34
7	1477.4	99.9	4615	12	US-10-190-435-36
8	1477.4	99.9	4702	12	US-10-190-435-38
9	1477.4	99.9	4716	12	US-10-190-435-17
10	1475.8	99.8	2742	12	US-10-190-435-20
11	1475.8	99.8	3930	12	US-10-190-435-9
12	1475.8	99.8	3930	12	US-10-190-435-10
13	1475.8	99.8	3930	12	US-10-190-435-11
14	1475.8	99.8	5145	12	US-10-190-435-12
15	1474.8	99.7	4713	12	US-10-190-435-59
16	1473.8	99.6	5184	12	US-10-190-435-58

17	1473.4	99.6	1479	12	US-09-967-464-67	Sequence 67, Appl
18	1472.4	99.6	2742	12	US-10-190-435-57	Sequence 57, Appl
19	1415.2	88.9	3531	12	US-10-190-435-13	Sequence 13, Appl
20	1415.2	88.9	3537	12	US-10-190-435-14	Sequence 14, Appl
21	1315.2	88.9	3537	12	US-10-190-435-15	Sequence 15, Appl
22	1275.2	86.2	1509	12	US-09-967-464-64	Sequence 64, Appl
23	1273.6	86.1	1509	12	US-09-967-464-68	Sequence 68, Appl
24	1196.2	80.9	2799	12	US-10-241-009-18	Sequence 18, Appl
25	1196.2	80.9	3205	12	US-10-241-009-17	Sequence 17, Appl
26	1196.2	80.9	3496	12	US-10-241-009-15	Sequence 15, Appl
27	1196.2	80.9	3999	12	US-10-241-009-9	Sequence 9, Appl
28	1196.2	80.9	3999	12	US-10-241-009-10	Sequence 10, Appl
29	1196.2	80.9	3999	12	US-10-241-009-11	Sequence 11, Appl
30	1196.2	80.9	4773	12	US-10-241-009-16	Sequence 16, Appl
31	1196.2	80.9	5274	12	US-10-241-009-12	Sequence 12, Appl
32	1195.2	80.8	2799	12	US-10-241-009-53	Sequence 53, Appl
33	1194.2	80.7	5283	12	US-10-241-009-54	Sequence 54, Appl
34	1193.5	80.7	4773	12	US-10-241-009-55	Sequence 55, Appl
35	1080.2	73.0	3564	12	US-10-241-009-13	Sequence 13, Appl
36	1080.2	73.0	3564	12	US-10-241-009-14	Sequence 14, Appl
37	976.6	66.0	1521	11	US-09-952-060-27	Sequence 27, Appl
38	976.6	66.0	1532	9	US-09-818-443-1	Sequence 1, Appl
39	976.6	66.0	1532	11	US-09-974-702-1	Sequence 1, Appl
40	976.6	66.0	37474	11	US-09-952-060-25	Sequence 25, Appl
41	973.6	65.8	4053	11	US-09-952-060-34	Sequence 34, Appl
42	949.4	64.2	4307	10	US-09-999-183-2	Sequence 2, Appl
43	933	63.1	1482	9	US-09-818-443-4	Sequence 4, Appl
44	930	62.9	1479	11	US-09-952-060-32	Sequence 32, Appl
45	916.2	61.9	3020	15	US-10-138-098-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1

US-09-967-464-63
; Sequence 63, Application US/09967464
; Publication No. US20030138453A1
; GENERAL INFORMATION:
; APPLICANT: O'Hagan, Derek
; APPLICANT: Otten, Gillis
; APPLICANT: Donnelly, John J.
; APPLICANT: Polo, John M.
; APPLICANT: Barnett, Susan
; APPLICANT: Singh, Mamohan
; APPLICANT: Ulmer, Jeffrey
; APPLICANT: Dubensky, Jr., Thomas W.
; TITLE OF INVENTION: MICROPARTICLES FOR DELIVERY OF HETEROLOGOUS NUCLEIC ACIDS
; FILE REFERENCE: PPI6269.004
; CURRENT APPLICATION NUMBER: US/09/967,464
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 60/236.105
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 60/315.905
; PRIOR FILING DATE: 2001-08-30
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 63
; LENGTH: 1479
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
US-09-967-464-63

Query Match 100.0% Score 1479; DB 12; Length 1479;
Best Local Similarity 100.0%; Pred. No. 2.9e-292;
Matches 1479; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGGGGCGCGCGCCAGCATCTGCGCGCGGGAAGCTGGACGCTGGAGCGCATCCGC	60
DB	1	ATGGGGCGCGCGCCAGCATCTGCGCGCGGGAAGCTGGACGCTGGAGCGCATCCGC	60
QY	61	CTGCGCGCGCGCGCGCAAGAGTGCTACATGATGAGGACCTGGTGGCGCCAGCGCGAG	120

Db 61 CTGGCCCCGGGGGGAAGAGTGTACATGATGAAGCAACCTGGTGTGGGCCAGCCGCGAG 120
QY 121 CTGGAGAAGTTGCGCCTGAACCCCGCCCTGCTGGAGACCCAGCGAGGGCTGCAAGCAGATC 180
Db 121 CTGGAGAAGTTGCGCCTGAACCCCGCCCTGCTGGAGACCCAGCGAGGGCTGCAAGCAGATC 180
QY 181 ATCCGCCAGCTGCACCCGCCCCCTGACAGACCGGCGAGGAGCTGAAGAGCCCTGTTCAC 240
Db 181 ATCCGCCAGCTGCACCCGCCCCCTGACAGACCGGCGAGGAGCTGAAGAGCCCTGTTCAC 240
QY 241 ACCGTGCCACCCCTGTACTCCGTGCAGAGAGAAGATCGAGTCCGGGACACCAAGAGGGCC 300
Db 241 ACCGTGCCACCCCTGTACTCCGTGCAGAGAGAAGATCGAGTCCGGGACACCAAGAGGGCC 300
QY 301 CTGGACAAGATCGAGGAGGAGCAGACAAGTGCACAGAGAAGATCCAGCAGGCGCAGGCC 360
Db 301 CTGGACAAGATCGAGGAGGAGCAGACAAGTGCACAGAGAAGATCCAGCAGGCGCAGGCC 360
QY 361 GCCGACRAGGGCAAGGTAGCCAGAACTACCCCATCTGTGCAGAACTGCAGGGCCAGATG 420
Db 361 GCCGACRAGGGCAAGGTAGCCAGAACTACCCCATCTGTGCAGAACTGCAGGGCCAGATG 420
QY 421 GTGCACCAAGGCCATCAGCCCCCGCCACCTGAACGGCTGGGTGAAGCTGATCGAGGAGAAG 480
Db 421 GTGCACCAAGGCCATCAGCCCCCGCCACCTGAACGGCTGGGTGAAGCTGATCGAGGAGAAG 480
QY 481 GCCTTCAGCCCCGAGGTGATCCCCATGTTACCGCCCTGAGCGAGGGGCGCCACCCGCGAG 540
Db 481 GCCTTCAGCCCCGAGGTGATCCCCATGTTACCGCCCTGAGCGAGGGGCGCCACCCGCGAG 540
QY 541 GACCTGAACACGATGTTGAACACCCGTGGGCGGCCACCGCCGCGCATGCAGATCTGAAG 600
Db 541 GACCTGAACACGATGTTGAACACCCGTGGGCGGCCACCGCCGCGCATGCAGATCTGAAG 600
QY 601 GACACCATCAACGAGGAGCGCGCGAGTGGACCCGCTGCACCCCTGCACCGCGCGCCGCC 660
Db 601 GACACCATCAACGAGGAGCGCGCGAGTGGACCCGCTGCACCCCTGCACCGCGCGCCGCC 660
QY 661 ATCGCCCCCGGCCAGATCGCGGAGCCCGCGCGGAGGACATCGCGGACCCACCAAGCCACC 720
Db 661 ATCGCCCCCGGCCAGATCGCGGAGCCCGCGCGGAGGACATCGCGGACCCACCAAGCCACC 720
QY 721 CTGCAGGAGCAGATCGCCTGGATGACCAACCAACCCCGCCCATCCCTGCGGAGATCTAC 780
Db 721 CTGCAGGAGCAGATCGCCTGGATGACCAACCAACCCCGCCCATCCCTGCGGAGATCTAC 780
QY 781 AAGCGGTGATCATCTTGGCCCTGAACAGATCGTGGGGAAGTACAGCCCTGTGAGCATC 840
Db 781 AAGCGGTGATCATCTTGGCCCTGAACAGATCGTGGGGAAGTACAGCCCTGTGAGCATC 840
QY 841 CTGGACATCAAGCAGGCGCCCAAGGAGCCCTTCGCGGACTACGTGACCCGCTCTTTCAG 900
Db 841 CTGGACATCAAGCAGGCGCCCAAGGAGCCCTTCGCGGACTACGTGACCCGCTCTTTCAG 900
QY 901 ACCCTGCGCGCGGAGCAGAGCACCCAGGAGGTGAAGAACTGGATGACCGACACCCCTGCTG 960
Db 901 ACCCTGCGCGCGGAGCAGAGCACCCAGGAGGTGAAGAACTGGATGACCGACACCCCTGCTG 960
QY 961 GTGCAGAACGCCCAACCCGACTGCAAGACCATCTCTGGCGGCTCTCGGCCCGCGGCGCAGC 1020
Db 961 GTGCAGAACGCCCAACCCGACTGCAAGACCATCTCTGGCGGCTCTCGGCCCGCGGCGCAGC 1020
QY 1021 CTGGAGGAGATGATGACCCCTGCCAGGCGTGGGCGGCCCCAGCCACAAAGGCCCGCGG 1080
Db 1021 CTGGAGGAGATGATGACCCCTGCCAGGCGTGGGCGGCCCCAGCCACAAAGGCCCGCGG 1080
QY 1081 CTGGCCGAGGCGATGAGCCAGGCCCAACACCAAGCGTGTATGATGCAGAGAGCAACTCAAG 1140
Db 1081 CTGGCCGAGGCGATGAGCCAGGCCCAACACCAAGCGTGTATGATGCAGAGAGCAACTCAAG 1140
QY 1141 GGCCCCCGCGCATCGTCAAGTCTTCAACTGCGGCAAGGAGGGCCACATCGCCCCCAAC 1200
Db 1141 GGCCCCCGCGCATCGTCAAGTCTTCAACTGCGGCAAGGAGGGCCACATCGCCCCCAAC 1200

QY 1201 TGCCGGCCCCCCCCCAGAGAGGGCTGCTGGAACTGGCGCAAGGAGCGCCACCAAGATGAAG 1260
Db 1201 TGCCGGCCCCCCCCCAGAGAGGGCTGCTGGAACTGGCGCAAGGAGCGCCACCAAGATGAAG 1260
QY 1261 GACTGCACCGAGCGCGCAGGCCAACTTCTCTGGCAAGATCTGGCCCGCCAGACAGGCGCCG 1320
Db 1261 GACTGCACCGAGCGCGCAGGCCAACTTCTCTGGCAAGATCTGGCCCGCCAGACAGGCGCCG 1320
QY 1321 CCGGGCAACTTCTCTGCAGAGCGCCCGCGAGCCCAACCGCCCCCGCCCGCGAGAGCTTCCGC 1380
Db 1321 CCGGGCAACTTCTCTGCAGAGCGCCCGCGAGCCCAACCGCCCCCGCCCGCGAGAGCTTCCGC 1380
QY 1381 TTCGAGGAGACCAACCCCGCGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440
Db 1381 TTCGAGGAGACCAACCCCGCGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440
QY 1441 CTGAAGAGCGCTGTTGCGCAACGACCCCGCTGAGCCAGTAA 1479
Db 1441 CTGAAGAGCGCTGTTGCGCAACGACCCCGCTGAGCCAGTAA 1479

RESULT 2
US-10-190-435-18
; Sequence 18, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEYER, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190.435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 3162
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GagNimul_C
US-10-190-435-18

Query Match 99.9%; Score 1477.4; DB 12; Length 3162;
Best Local Similarity 99.9%; Pred. NO. 5.9e-252;
Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGGGCGCGCGCGCCAGCATCTCGCGCGCGGCAAGCTGACCGCTGGGAGCGCAICCGC 60
Db 7 ATGGGGCGCGCGCGCCAGCATCTCGCGCGCGGCAAGCTGACCGCTGGGAGCGCATCCGC 66
QY 61 CTGCGCCCCCGCGCGGCAAGAGTGTCTACATGATGAAGCACTGGTGTGGGCCAGCGCGAG 120
Db 67 CTGCGCCCCCGCGCGGCAAGAGTGTCTACATGATGAAGCACTGGTGTGGGCCAGCGCGAG 126
QY 121 CTGGAGAAGTTGCGCCTGAACCCCGCCCTGCTGGAGACCGAGGAGGCTGCAAGCAGATC 180
Db 127 CTGGAGAAGTTGCGCCTGAACCCCGCCCTGCTGGAGACCGAGGAGGCTGAAGAGCCTGTTCAAC 186
QY 181 ATCCGCCAGCTGCACCCCGCCCTGCAGACCGGCGAGGAGGAGCTGAAGAGCCTGTTCAC 240
Db 187 ATCCGCCAGCTGCACCCCGCCCTGCAGACCGGCGAGGAGGAGCTGAAGAGCCTGTTCAC 246
QY 241 ACCGTGCCACCCCTGTACTCGTGCACAGAGAAGATCGAGTCCGGGACACCAAGAGGCC 300
Db 247 ACCGTGCCACCCCTGTACTCGTGCACAGAGAAGATCGAGTCCGGGACACCAAGAGGCC 306
QY 301 CTGGACAAGATCGAGGAGGAGCAGACAAGTGCACAGAGAAGATCCAGCAGGCGCCAGGCC 360
Db 301 CTGGACAAGATCGAGGAGGAGCAGACAAGTGCACAGAGAAGATCCAGCAGGCGCCAGGCC 360

Db 307 CTGGACAAGATCGAGGAGGACAGAAAGTGTCCAGAGAGAGATCCAGACGCCCGGAGGGCC 366
QY 361 GCGGACAAGGCGAAGGTGAGCCAGAACTACCCCACTGTCAGAGAACTGCGAGGGCGAGATG 420
Db 367 GCGGACAAGGCGAAGGTGAGCCAGAACTACCCCACTGTCAGAGAACTGCGAGGGCGAGATG 426
QY 421 GTGCACACAGCCATCAGCCCGCCGACCCCTGAACCCCTGGGTGAAGGTGATCGAGGAGAG 480
Db 427 GTGCACACAGCCATCAGCCCGCCGACCCCTGAACCCCTGGGTGAAGGTGATCGAGGAGAG 486
QY 481 GCCTTCAGCCCGGAGGTGATCCCCATGTCACCCCGCCGACCCCTGAACCCCTGGGTGAAGGTGATCGAGGAGAG 540
Db 487 GCCTTCAGCCCGGAGGTGATCCCCATGTCACCCCGCCGACCCCTGAACCCCTGGGTGAAGGTGATCGAGGAGAG 546
QY 541 GACCTGAACACGATGTTGAACACCGTGGCGGGCCACAGCCCGCCATGTCAGATGCTGAAG 600
Db 547 GACCTGAACACGATGTTGAACACCGTGGCGGGCCACAGCCCGCCATGTCAGATGCTGAAG 606
QY 601 GACACCATCAACGAGGAGGCGCCGAGTGGGAGCCGCTGCACCCCGCTGCACGCGCGGCGCC 660
Db 607 GACACCATCAACGAGGAGGCGCCGAGTGGGAGCCGCTGCACCCCGCTGCACGCGCGGCGCC 666
QY 661 ATCGCCCGCCGCGCAGATGCGCGGAGCCCGCGGAGCGACATCGCGGACCCACACGAGACC 720
Db 667 ATCGCCCGCCGCGCAGATGCGCGGAGCCCGCGGAGCGACATCGCGGACCCACACGAGACC 726
QY 721 CTGCAGGAGCAGATCGCCTGCGATGACCAAGACACCCCGCATCCCGTGGGCGGACATCTAC 780
Db 727 CTGCAGGAGCAGATCGCCTGCGATGACCAAGACACCCCGCATCCCGTGGGCGGACATCTAC 786
QY 781 AAGCGGTGATCATCCTGGGCCCTGAACAGAGATCGTGCAGATGATACAGCCCGCGTGAGCAAC 840
Db 787 AAGCGGTGATCATCCTGGGCCCTGAACAGAGATCGTGCAGATGATACAGCCCGCGTGAGCAATC 846
QY 841 CTGGACATCAAGCAGGCGCCCAAGGAGCCCTTCCCGGACTACGTGGACCGCTCTCTCAAG 900
Db 847 CTGGACATCAAGCAGGCGCCCAAGGAGCCCTTCCCGGACTACGTGGACCGCTCTCTCAAG 906
QY 901 ACCCTGCGCGCGGAGCAGACACCCAGGAGGTGAAGAACTGGATGACCGACACCCCTGCTG 960
Db 907 ACCCTGCGCGCGGAGCAGACACCCAGGAGGTGAAGAACTGGATGACCGACACCCCTGCTG 966
QY 961 GTGCAGAACGCGCAACCCCGACTGCAAGACCATCTCTCGCGCTCTCGGCGCCCGCGGAGACC 1020
Db 967 GTGCAGAACGCGCAACCCCGACTGCAAGACCATCTCTCGCGCTCTCGGCGCCCGCGGAGACC 1026
QY 1021 CTGGAGGAGATGATGACCGCTGCCAGGGCGGTGGGGGCGCCACAGCCAAAGGCGCGCGTG 1080
Db 1027 CTGGAGGAGATGATGACCGCTGCCAGGGCGGTGGGGGCGCCACAGCCAAAGGCGCGCGTG 1086
QY 1081 CTGGCGGAGGATGAGCCAGGCGCCAAACACAGCGGTGATGTCAGAGAGAGCAACTTCAAG 1140
Db 1087 CTGGCGGAGGATGAGCCAGGCGCCAAACACAGCGGTGATGTCAGAGAGAGCAACTTCAAG 1146
QY 1141 GCGCCCGCGCGCATCGTCAAGTGTCTCAACTGCGCGCAAGGAGGGCCACATCGCCCGGCAAC 1200
Db 1147 GCGCCCGCGCGCATCGTCAAGTGTCTCAACTGCGCGCAAGGAGGGCCACATCGCCCGGCAAC 1206
QY 1201 TCGCGCGCGCGCGCAAGAGGGCTGTGGAAGTGGCGAAGGAGGGGCGACACAGATGAAG 1260
Db 1207 TCGCGCGCGCGCGCAAGAGGGCTGTGGAAGTGGCGAAGGAGGGGCGACACAGATGAAG 1266
QY 1261 GACTGCACCGAGCGCGCAAACTTCTTGGGCAAGATCTGGCCCGAGGCGACACAAGGGGCGG 1320
Db 1267 GACTGCACCGAGCGCGCAAACTTCTTGGGCAAGATCTGGCCCGAGGCGACACAAGGGGCGG 1326
QY 1321 CCGGCAACTTCTTGCAGAGCGCGCGCAACCGCGCCACCGCCCGCGCGAGAGCGTTCGCG 1380
Db 1327 CCGGCAACTTCTTGCAGAGCGCGCGCAACCGCGCCACCGCCCGCGCGAGAGCTTCGCG 1386
QY 1381 TTCGAGGAGACACCCCGCGCGAGAGAGCAGGAGAGCAAGGACCGCGAGACCCCTGACCCAGC 1440
Db 1387 TTCGAGGAGACACCCCGCGCGAGAGAGCAGGAGAGCAAGGACCGCGAGACCCCTGACCCAGC 1446

QY 1441 CTGAAGAGCCGTGTCGGCAACGACCCCGCTGAGCCAGTAA 1479
Db 1447 CTGAAGAGCCGTGTCGGCAACGACCCCGCTGAGCCAGAAA 1485

RESULT 3

US-10-190-435-16
; Sequence 16, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEDEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PPI8133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190.435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 3462
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GagProtInaRimul_C
US-10-190-435-16

Query Match 99.9%; Score 1477.4; DB 12; Length 3462;
Best Local Similarity 99.9%; Pred. No. 5.9e-292;
Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGGCGCCGCCAGCATCTCGCGGGCGCAAGCTGGACGCCCTGGGAGCGCATCCGC 60
Db 7 ATGGGGCGCCGCCAGCATCTCGCGGGCGCAAGCTGGACGCCCTGGGAGCGCATCCGC 66
QY 61 CTGCGCCCGCGGGCAAGAGTGTACATGATGAAGCACTGGTGTGGCCAGCCGCGAG 120
Db 67 CTGCGCCCGCGGGCAAGAGTGTACATGATGAAGCACTGGTGTGGCCAGCCGCGAG 126
QY 121 CTGGAGAAGTTCGCCCTGAACCCCGCGCTGCTGGAGACGAGGGCTGCAAGCAGATC 180
Db 127 CTGGAGAAGTTCGCCCTGAACCCCGCGCTGCTGGAGACGAGGGCTGCAAGCAGATC 186
QY 181 ATCGCGCAGCTGCACCCCGCTGCAGACCGCGCAGCGAGAGCTGAAGACCTGTTCAAC 240
Db 187 ATCGCGCAGCTGCACCCCGCTGCAGACCGCGCAGCGAGAGCTGAAGACCTGTTCAAC 246
QY 241 ACCGTGGCCACCCCTGTACTGCGTGCACGAGAAGATCGAGGTCCCGACACCAAGGAGCC 300
Db 247 ACCGTGGCCACCCCTGTACTGCGTGCACGAGAAGATCGAGGTCCCGACACCAAGGAGCC 306
QY 301 CTGGACAAGATCGAGGAGGAGGAGCAAGAGTCCCGCAGCAGAGATCCAGGAGCGCGAGCC 360
Db 307 CTGGACAAGATCGAGGAGGAGGAGCAAGAGTCCCGCAGCAGAGATCCAGGAGCGCGAGCC 366
QY 361 GCGCACAAAGGCAAGGTGAGCCAGAACTACCCCATCTGTGAGAACTCCAGGGCCAGATG 420
Db 367 GCGCACAAAGGCAAGGTGAGCCAGAACTACCCCATCTGTGAGAACTCCAGGGCCAGATG 426
QY 421 GTGCACACAGGCGCATCAGCCCGCGCACCCCTGAACCCCTGGGTGAAGGTGATCGAGGAGAG 480
Db 427 GTGCACACAGGCGCATCAGCCCGCGCACCCCTGAACCCCTGGGTGAAGGTGATCGAGGAGAG 486
QY 481 GCCTTCAGCCCGGAGGTGATCCCCATGTCACCCCGCCCTGACCGGAGGCGCCACCCCGAG 540
Db 487 GCCTTCAGCCCGGAGGTGATCCCCATGTCACCCCGCCCTGACCGGAGGCGCCACCCCGAG 546
QY 541 GACCTGAACACGATGTTGAACACCGTGGCGGGCCACAGCCCGCCATGTCAGATGCTGAAG 600
Db 547 GACCTGAACACGATGTTGAACACCGTGGCGGGCCACAGCCCGCCATGTCAGATGCTGAAG 606


```
Db 547 GACCIGAACACGATGTTGACACCGTGGCGCGCCACACAGCGCCGATCCAGATGCTGAAG 606
QY 601 GACACCATCAACAGGAGGCGCGCGAGTGGACCGCGTGCACCCCGTGCACGCGCGCC 660
Db 607 GACACCATCAACAGGAGGCGCGCGAGTGGACCGCGTGCACCCCGTGCACGCGCGCC 666
QY 661 ATCGCCCCCGCCAGATGCGCGAGCCCGCGGCGAGCGACATCGCCGGACCAACAGCACC 720
Db 667 ATCGCCCCCGCCAGATGCGCGAGCCCGCGGCGAGCGACATCGCCGGACCAACAGCACC 726
QY 721 CTGCAGGAGCAGATCGCGCTGGATGACCAAGCAACCCCGCATCCCGTGGCGACATCTAC 780
Db 727 CTGCAGGAGCAGATCGCGCTGGATGACCAAGCAACCCCGCATCCCGTGGCGACATCTAC 786
QY 781 AAGCGGTGGATCATCTGGGCTGAACAGATCGTGGGATGTACAGCCCGGTGAGCATC 840
Db 787 AAGCGGTGGATCATCTGGGCTGAACAGATCGTGGGATGTACAGCCCGGTGAGCATC 846
QY 841 CTGGACATCAAGCAGGCGCCCAAGGAGCCCTTCCGCGACTACGTGGACCCCTCTTCAAG 900
Db 847 CTGGACATCAAGCAGGCGCCCAAGGAGCCCTTCCGCGACTACGTGGACCCCTCTTCAAG 906
QY 901 ACCCTGCGCGCGAGCAGACACCCAGGAGGIGAGAACTGGATGACCGACACCCCTGCTG 960
Db 907 ACCCTGCGCGCGAGCAGACACCCAGGAGGIGAGAACTGGATGACCGACACCCCTGCTG 966
QY 961 GTGCAGAACGCCAACCCGACTGCAAGACCATCTCGCGGCTCTCGGCGCGCGGCGAGC 1020
Db 967 GTGCAGAACGCCAACCCGACTGCAAGACCATCTCGCGGCTCTCGGCGCGCGGCGAGC 1026
QY 1021 CTGCAGGAGATGATGACCGCTGCCAGGCGGTGGCGGCGCCAGCCCAAGGCGCGGCTG 1080
Db 1027 CTGCAGGAGATGATGACCGCTGCCAGGCGGTGGCGGCGCCAGCCCAAGGCGCGGCTG 1086
QY 1081 CTGGCGGAGGCGATGAGCCAGGCCCAACACAGCGTGTATGATGACGAGAGCAACTTCAAG 1140
Db 1087 CTGGCGGAGGCGATGAGCCAGGCCCAACACAGCGTGTATGATGACGAGAGCAACTTCAAG 1146
QY 1141 GCGCCCGCGCGATCGTCAAGTGTCTCAACTGCGGCAAGGAGGCGCCACATCGCGCGCAAG 1200
Db 1147 GCGCCCGCGCGATCGTCAAGTGTCTCAACTGCGGCAAGGAGGCGCCACATCGCGCGCAAG 1206
QY 1201 TGCGCGCGCCCGCCNAGAGGGCTGTGGAAGTGGGCAAGGAGGCGCCACAGATCAAG 1260
Db 1207 TGCGCGCGCCCGCCNAGAGGGCTGTGGAAGTGGGCAAGGAGGCGCCACAGATCAAG 1266
QY 1261 GACTGCACCGAGCGCCAGCGCAACTTCTGGCAAGATCTGGCCAGCGCCACAGAGCGCGC 1320
Db 1267 GACTGCACCGAGCGCCAGCGCAACTTCTGGCAAGATCTGGCCAGCGCCACAGAGCGCGC 1326
QY 1321 CCGCGCAACTTCTGTCAGAGCGCGCCCGAGCGCCAGCGCCCGCGCGAGAGCTTGGC 1380
Db 1327 CCGCGCAACTTCTGTCAGAGCGCGCCCGAGCGCCAGCGCCCGCGCGAGAGCTTGGC 1386
QY 1381 TTCGAGGAGACACCCCGCGCAGAGCAGAGCAGAGCAAGACCGCGAGACCCCTGACCAAG 1440
Db 1387 TTCGAGGAGACACCCCGCGCAGAGCAGAGCAGAGCAGAGCAGAGCAGAGCAGAGCAG 1446
QY 1441 CTGAAGAGCGCTGTTGCGCAACGACCCCTTGAGCCAGTAA 1479
Db 1447 CTGAAGAGCGCTGTTGCGCAACGACCCCTTGAGCCAGAAA 1485
```

RESULT 4

```
US-10-190-435-19
; Sequence 19, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGERE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susar
; APPLICANT: VAN RENSBURG, Estrelita J.
```

```
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 4419
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GagRImutTatRevNef_C
US-10-190-435-19
```

```
Query Match 99.9%; Score 1477.4; DB 12; Length 4419;
Best Local Similarity 99.9%; Pred. No. 5.8e-292;
Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 ATGGGCGCGCGCGCCAGCATCCTGCGGCGCGGCAAGCTGGACGCTGGAGCGCATCCGC 60
Db 7 ATGGGCGCGCGCGCCAGCATCCTGCGGCGCGGCAAGCTGGACGCTGGAGCGCATCCGC 66
QY 61 CTGCGCGCGCGCGCGCAAGAGTGTCTACATGATGAAGCACTGGTGTGGGCGCAGCGCGAG 120
Db 67 CTGCGCGCGCGCGCGCAAGAGTGTCTACATGATGAAGCACTGGTGTGGGCGCAGCGCGAG 126
QY 121 CTGGAGAAAGTTCGCCCTGAACCCCGCGCTGCTGGAGACAGCGAGGGCTCCANGCAGATC 180
Db 127 CTGGAGAAAGTTCGCCCTGAACCCCGCGCTGCTGGAGACAGCGAGGGCTCCANGCAGATC 186
QY 181 ATCCGCCAGCTGCACCCCGCCCTGCAGACCGGCGAGCGAGAGCTGAAGAGCCCTGTTCAAC 240
Db 187 ATCCGCCAGCTGCACCCCGCCCTGCAGACCGGCGAGCGAGAGCTGAAGAGCCCTGTTCAAC 246
QY 241 ACCGTGGCGCACCCCTGTACTGCGTGCACGAGAGATCGAGGTCCGCGACACCAAGGAGGCC 300
Db 247 ACCGTGGCGCACCCCTGTACTGCGTGCACGAGAGATCGAGGTCCGCGACACCAAGGAGGCC 306
QY 301 CTGGACAAGATCGAGGAGGAGGAGCAACAGTGCAGAGAGATCCAGCAGGCGCGAGGCC 360
Db 307 CTGGACAAGATCGAGGAGGAGGAGCAACAGTGCAGAGAGATCCAGCAGGCGCGAGGCC 366
QY 361 GCGGACAAGGCAAGGTGAGCCCGCAACTAGCCCATCTGTGCAAACTGCGAGGCGCCAGATG 420
Db 367 GCGGACAAGGCAAGGTGAGCCCGCAACTAGCCCATCTGTGCAAACTGCGAGGCGCCAGATG 426
QY 421 GCGCACCGAGCCCATCAGCCCGCGCACCTGAAAGCGCTGGGTGAAGTGAAGGAGGAAG 480
Db 427 GCGCACCGAGCCCATCAGCCCGCGCACCTGAAAGCGCTGGGTGAAGTGAAGGAGGAAG 486
QY 481 GCGTTCAGCCCGAGGTGATCCCGCATGTTACCGCCCGTGAAGCGGAGGCGCGCCACCCCGAG 540
Db 487 GCGTTCAGCCCGAGGTGATCCCGCATGTTACCGCCCGTGAAGCGGAGGCGCGCCACCCCGAG 546
QY 541 GACCTGAACACGATGTTGAACACCGTGGGCGCGCCACCGAGCCCGCATGCAGATGCTGAAG 600
Db 547 GACCTGAACACGATGTTGAACACCGTGGGCGCGCCACCGAGCCCGCATGCAGATGCTGAAG 606
QY 601 GACACCATCAACGAGGAGGCGCGCGAGTGGGACCGCGTGCACCCCGTGCACCGCGCGGCC 660
Db 607 GACACCATCAACGAGGAGGCGCGCGAGTGGGACCGCGTGCACCCCGTGCACCGCGCGGCC 666
QY 661 ATCGCCCCCGCCAGATGCGCGAGCCCGCGGCGAGCGACATCGCGGCGACCAACAGCACC 720
Db 667 ATCGCCCCCGCCAGATGCGCGAGCCCGCGGCGAGCGACATCGCGGCGACCAACAGCACC 726
QY 721 CTGCAGGAGCAGATCGCGCTGGATGACCAAGCAACCCCGCATCCCGTGGCGGCGACATCTAC 780
Db 727 CTGCAGGAGCAGATCGCGCTGGATGACCAAGCAACCCCGCATCCCGTGGCGGCGACATCTAC 786
QY 781 AAGCGGTGGATCATCTGGGCTGAACAGATCGTGGGATGTACAGCCCGGTGAGCATC 840
Db 1447 CTGAAGAGCGCTGTTGCGCAACGACCCCTTGAGCCAGAAA 1485
```

Db 787 AAGCGGTGGAATCATCCCTGGGCTTGAAACAAGATCTGCGGATGTACAGCCCGGTGAGCAATC 815
QY 841 CTGGACATCAAGCAGGCGCCCAAGGAGCCCTTCGGGAGTACGTGACCGCTTCTTCAAG 909
Db 847 CTGGACATCAAGCAGGCGCCCAAGGAGCCCTTCGGGAGTACGTGACCGCTTCTTCAAG 906
QY 901 ACCCTGCGGCGGAGCAGACAGCACCCAGAGGTGAAGAACTGGATGACCGAGCACCGTCTG 950
Db 907 ACCCTGCGGCGGAGCAGACAGCACCCAGAGGTGAAGAACTGGATGACCGAGCACCGTCTG 946
QY 961 GTGCAGAAAGCCCAACCTTGACTGCAAGAGCATCTCTGAGGCTCTGCGCCCGGCGAGG 1020
Db 967 GTGCAGAAAGCCCAACCCGACTGCAAGAGCATCTCTGAGGCTCTGCGCCCGGCGAGG 1026
QY 1021 CTGGAGGAGATGATGACCGCCCTGCCAGGCGCTGGGCGGCCCAAGCAAGAGGCGGAGG 1060
Db 1027 CTGGAGGAGATGATGACCGCCCTGCCAGGCGCTGGGCGGCCCAAGCAAGAGGCGGAGG 1056
QY 1081 CTGGCCGAGCGGATGAGCCAGGCGCAACACAGCGGTGATGATGCAAGAGCAAGCTTCAAG 1110
Db 1087 CTGGCCGAGCGGATGAGCCAGGCGCAACACCGCGGTGATGATGCAAGAGCAAGCTTCAAG 1116
QY 1141 GGCCCCCGGCGCATCTCAAGTGCTCAACTGCGGCAAGGAGGGGCGACATCGGCGCGCAAC 1200
Db 1147 GGCCCCCGGCGCATCTCAAGTGCTCAACTGCGGCAAGGAGGGGCGACATCGGCGCGCAAC 1206
QY 1201 TGCCGCGCCCGCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGGAGGGCCACACAGATGAAG 1250
Db 1207 TGCCGCGCCCGCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGGAGGGCCACACAGATGAAG 1256
QY 1261 GACTGCACCGAGCGCCAGGCCCAACTTCTTGGCAAGATCTGGCCAGCCACACAGGCGCGC 1320
Db 1267 GACTGCACCGAGCGCCAGGCCCAACTTCTTGGCAAGATCTGGCCAGCCACACAGGCGCGC 1326
QY 1321 CCCGGCAACTTCCCTGCGAGAGCGCGCCGAGCGCCACCGCCCGCCCGCGAGAGCTTCCGC 1380
Db 1327 CCCGGCAACTTCCCTGCGAGAGCGCGCCGAGCGCCACCGCCCGCCCGCGAGAGCTTCCGC 1386
QY 1381 TTCGAGGAGACCAACCCCGCGCCAGAGCAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440
Db 1387 TTCGAGGAGACCAACCCCGCGCCAGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1446
QY 1441 CTGAAGAGCGCTGTTCCGCAACGACCCCGCTGAGCCAGTAA 1479
Db 1447 CTGAAGAGCGCTGTTCCGCAACGACCCCGCTGAGCCAGTAA 1485

RESULT 5

US-10-190-435-35
; Sequence 35, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEYER, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PPI8133.003 / 2302-1833
; CURRENT APPLICATION NUMBER: US/10/190.435
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 35
; LENGTH: 4483
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: gp160mod.rv1.dv1v2-gagmod.Bw965

US-10-190-435-35

Query Match 99.9%; Score 1477.4; DB 12; Length 4483;
Best Local Similarity 99.9%; Pred. No. 5.8e-292;
Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGGCGCGCGCGGCGGAGCATCTCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 50
Db 3005 ATGGGCGCGCGCGGCGGAGCATCTCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3064
QY 61 CTGGCGCGCGCGGCGGCAAGTGTACATGATGATGATGATGATGATGATGATGATGATGATGATG 120
Db 3065 CTGGCGCGCGCGGCGGCAAGTGTACATGATGATGATGATGATGATGATGATGATGATGATG 3124
QY 121 CTGGAGAAAGTTCGCGCTGAACCCCGCGCTGCTGGAGAGGAGGAGGAGGAGGAGGAGGAGG 180
Db 3125 CTGGAGAAAGTTCGCGCTGAACCCCGCGCTGCTGGAGAGGAGGAGGAGGAGGAGGAGGAGG 3184
QY 181 ATCCGCGAGCTGCACCGCGCGCTGCAGACCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240
Db 3185 ATCCGCGAGCTGCACCGCGCGCTGCAGACCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3244
QY 241 ACCGTGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
Db 3245 ACCGTGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3304
QY 301 CTGGACAAGATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
Db 3305 CTGGACAAGATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3364
QY 361 GCGGACAAGGCGCAAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420
Db 3365 GCGGACAAGGCGCAAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3424
QY 421 GTGCACGAGGCGCATCAGCG 480
Db 3425 GTGCACGAGGCGCATCAGCG 3484
QY 481 GCCTTCAGCG 540
Db 3485 GCCTTCAGCG 3544
QY 541 GACCTGAAACAGCATGTTGAACACCCCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
Db 3545 GACCTGAAACAGCATGTTGAACACCCCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3604
QY 601 GACACCATCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
Db 3605 GACACCATCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3664
QY 661 ATCGCGCGCGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
Db 3665 ATCGCGCGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3724
QY 721 CTGCAGGAGCAGATCGCTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
Db 3725 CTGCAGGAGCAGATCGCTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3784
QY 781 AAGCGGTGGATCATCTGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
Db 3785 AAGCGGTGGATCATCTGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 3844
QY 841 CTGGACATCAAGCAGGCG 900
Db 3845 CTGGACATCAAGCAGGCG 3904
QY 901 ACCCTGCGCGCGCGGAGCAGACCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960
Db 3905 ACCCTGCGCGCGCGGAGCAGACCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3964
QY 961 GTGCAGAAAGCG 1020
Db 3965 GTGCAGAAAGCG 4024
QY 1021 CTGGAGGAGATGATGACCGCGCTGCCAGGCGGCTGGGCGGCGGCGGCGGCGGCGGCGGCGG 1080

```
Db 4025 CTGGAGAGATGATGACGGCTGCCAGGCGCTGGGGGCCCCAGCCACAAAGCGCGCGG 4084
QY 1081 CTGGCCGAGCGAIGAGCCAGGCCAACACACACCGTGAATGATGCGAAGAGCAATTCACG 1140
Db 4085 CTGGCCGAGCGAIGAGCCAGGCCAACACACACCGTGAATGATGCGAAGAGCAATTCACG 4144
QY 1141 GGCCCCCGCGCATCGTCAAGTGTCTCAACTGCGGCAAGGAGGGCCACATCGCCCGCAAC 1200
Db 4145 GGCCCCCGCGCATCGTCAAGTGTCTCAACTGCGGCAAGGAGGGCCACATCGCCCGCAAC 4204
QY 1201 TGCCCGCGCCCCCGCAAGAGGGCTGTGGAAGTGGGCAAGGAGGGCCACCATGATGAAG 1260
Db 4205 TGCCCGCGCCCCCGCAAGAGGGCTGTGGAAGTGGGCAAGGAGGGCCACCATGATGAAG 4264
QY 1261 GACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCAGCCACAAAGCGCGC 1320
Db 4265 GACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCAGCCACAAAGCGCGC 4324
QY 1321 CCGGCAACTTCCTGCAGAGCGCCCGCGAGCCGACCGCCCCCGCGGAGAGCTCCGC 1380
Db 4325 CCGGCAACTTCCTGCAGAGCGCCCGCGAGCCGACCGCCCCCGCGGAGAGCTCCGC 4384
QY 1381 TTCGAGAGACCAACCCCGCGCCAGAGAGGAGAGCAAGGACCGCGGAGACCCCTGACCAAG 1440
Db 4385 TTCGAGAGACCAACCCCGCGCCAGAGAGGAGAGCAAGGACCGCGGAGACCCCTGACCAAG 4444
QY 1441 CTGAAGAGCGCTTCGGCAACGACCCCGCTGAGCCAA 1479
Db 4445 CTGAAGAGCGCTTCGGCAACGACCCCGCTGAGCCAA 4483
```

RESULT 6

```
US-10-190-435-34
; Sequence 34, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: P18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190.435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34
; LENGTH: 4606
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: gpl60mod.TV1.dv1-gagmod.RW965
```

```
US-10-190-435-34
Query Match 99.9%; Score 1477.4; Db 12; Length 4606;
Best Local Similarity 99.9%; Pred. No. 5.8e-292;
Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGGCGCGCGCCAGCATCTCGCGGGCGGCAAGTGGACCGCTGGGAGCGCATCCGC 60
Db 3128 ATGGGGCGCGCGCCAGCATCTCGCGGGCGGCAAGTGGACCGCTGGGAGCGCATCCGC 3187
QY 61 CTGGCCCCCGCGGCAAGAGTGTCTACATGATGAAGACCTGGTGTGGGCCAGCGCGAG 120
Db 3188 CTGGCCCCCGCGGCAAGAGTGTCTACATGATGAAGACCTGGTGTGGGCCAGCGCGAG 3247
QY 121 CTGGAGAACTTCGCCCTGAACCCCGCGCTGTGGAGACCAGCGAGGGCTGCAAGAGATC 190
Db 3248 CTGGAGAACTTCGCCCTGAACCCCGCGCTGTGGAGACCAGCGAGGGCTGCAAGAGATC 3307
```

```
QY 181 ATCCGCCAGCTGCACCCCGCCCTGCAGACCGGCGAGGAGGCTGAGAGCCCTGTTCAAC 240
Db 3308 ATCCGCCAGCTGCACCCCGCCCTGCAGACCGGCGAGGAGGCTGAGAGCCCTGTTCAAC 3367
QY 241 ACCGIGGCCACCCCTGTACTGCGTGCACGAGAAGATCCAGGTCGCGCACACCAAGGAGGCC 300
Db 3368 ACCGIGGCCACCCCTGTACTGCGTGCACGAGAAGATCCAGGTCGCGCACACCAAGGAGGCC 3427
QY 301 CTGGACACAGATCGAGGAGGAGCAGAACAAAGTGCCAGCAGAGAATCCAGCAGGCCGAGGCC 360
Db 3428 CTGGACACAGATCGAGGAGGAGCAGAACAAAGTGCCAGCAGAGAATCCAGCAGGCCGAGGCC 3487
QY 361 GCCGACACAGGCAAGGTGAGCCAGAACTACCCCATCTGTCAGAACTGTCAGGGGCCAGATG 420
Db 3488 GCCGACACAGGCAAGGTGAGCCAGAACTACCCCATCTGTCAGAACTGTCAGGGGCCAGATG 3547
QY 421 GTGCACACAGGCCATCAGCCCCCGCACCCCTGAACCGCTGGGTGAAGGTGATCGAGGAGAAG 480
Db 3548 GTGCACACAGGCCATCAGCCCCCGCACCCCTGAACCGCTGGGTGAAGGTGATCGAGGAGAAG 3607
QY 481 GCCTTCAGCCCCGAGGTGATCCCATGTTQACCGCCCTGAGCGAGGCGCCACCCCGCAG 540
Db 3608 GCCTTCAGCCCCGAGGTGATCCCATGTTQACCGCCCTGAGCGAGGCGCCACCCCGCAG 3667
QY 541 GACCTGAACACAGATGTTGAACACCGCTGGCGGCGCACAGGCCCGCATGCGATGCTGAAG 600
Db 3668 GACCTGAACACAGATGTTGAACACCGCTGGCGGCGCACAGGCCCGCATGCGATGCTGAAG 3727
QY 601 GACACCATCAACAGGAGGCGCGCGGAGCGGCGGAGCGACATCGCGGCGCACACCCAGCAC 660
Db 3728 GACACCATCAACAGGAGGCGCGCGGAGCGGAGCGGAGCGGCGGCGGCGGCGGCGGCGG 3787
QY 661 ATCGCCCCCGCGCAGATGCGCGGAGCGCGCGGCGGAGCGACATCGCGGCGCACACCCAGCAC 720
Db 3788 ATCGCCCCCGCGCAGATGCGCGGAGCGCGCGGCGGAGCGACATCGCGGCGCACACCCAGCAC 3847
QY 721 CTGCAGGAGCAGATCGCCTGGATGACCAAGATCGTGGGATGTACAGCCCGGTGAGCATC 780
Db 3848 CTGCAGGAGCAGATCGCCTGGATGACCAAGATCGTGGGATGTACAGCCCGGTGAGCATC 3907
QY 781 AAGCGGTGATCATCTCGGCGCTGAACAAAGATCGTGGGATGTACAGCCCGGTGAGCATC 840
Db 3908 AAGCGGTGATCATCTCGGCGCTGAACAAAGATCGTGGGATGTACAGCCCGGTGAGCATC 3967
QY 841 CTGCACATCAAGCAGGCGCCCAAGGAGCGCTTCGCGGACTACGTGGAGCGCTTCTTCAAG 900
Db 3968 CTGGACA:CAAGCAGGCGCCCAAGGAGCGCTTCGCGGACTACGTGGAGCGCTTCTTCAAG 4027
QY 901 ACCCTGGCGCGCAGGACAGCAACCCAGGAGGTGAAGAACTGGATGACCGGACACCCCTGCTG 960
Db 4028 ACCCTGGCGCGCAGGACAGCAACCCAGGAGGTGAAGAACTGGATGACCGGACACCCCTGCTG 4087
QY 961 GTGCAGAACCGCAACCCCGACTTCAAGACCATCTCGCGGCTCTCGGCGCGCGGCGGCGAGC 1020
Db 4088 GTGCAGAACCGCAACCCCGACTTCAAGACCATCTCGCGGCTCTCGGCGCGCGGCGGCGAGC 4147
QY 1021 CTGGAGGAGATGATGACCCCTGCCAGGGTGTGGCGGCGCCAGCCACAGGCCCGCGGTG 1080
Db 4148 CTGGAGGAGATGATGACCCCTGCCAGGGTGTGGCGGCGCCAGCCACAGGCCCGCGGTG 4207
QY 1081 CTGGCCGAGCGGATGAGCCAGGCCAACACAGCGTGTATGATGCAGAGAGCAACTTCAAG 1140
Db 4208 CTGGCCGAGCGGATGAGCCAGGCCAACACAGCGTGTATGATGCAGAGAGCAACTTCAAG 4267
QY 1141 GGCCCCCGCGCATCGTCAAGTGTCTCAACTGCGGCAAGGAGGGCCACATCGCCCGCAAC 1200
Db 4268 GGCCCCCGCGCATCGTCAAGTGTCTCAACTGCGGCAAGGAGGGCCACATCGCCCGCAAC 4327
QY 1201 TGCCCGCGCGCGCGCAAGAGGGCTGTGGAAGTGGCGCAAGGAGGGCCACCATGATGAAG 1260
Db 4328 TGCCCGCGCGCGCGCAAGAGGGCTGTGGAAGTGGCGCAAGGAGGGCCACCATGATGAAG 4387
```


; CURRENT APPLICATION NUMBER: US/10/190.435	
; CURRENT FILING DATE: 2002-12-30	
; NUMBER OF SEQ ID NOS: 319	
; SOFTWARE: PatentIn Ver. 2.0	
; SEQ ID NO 17	
; LENGTH: 4716	
; TYPE: DNA	
; ORGANISM: Artificial Sequence	
; FEATURE:	
; OTHER INFORMATION: Description of Artificial Sequence:	
; OTHER INFORMATION: GagProtInArtMutTatRevNef_C	
US-10-190-435-17	
Query Match 99.9%; Score 1477.4; DB 12; Length 4716;	
Best Local Similarity 99.9%; Pred. No. 5.8e-292;	
Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
QY 1	ATGGGGCGCGCGCCAGCATCCCTGGGGCGCGCGCAAGCTGGAGCGCTGGGAGCGCATCCGC 60
DB	
QY 7	ATGGGGCGCGCGCGCCAGCATCCCTGGGGCGCGCGCAAGCTGGAGCGCTGGGAGCGCATCCGC 66
DB	
QY 61	CTGGCGCGCGCGCGCGCAAGAGTGTACATGATGAAGCAACCTGGTGGGGCGCGCGAG 120
DB	
QY 67	CTGGCGCGCGCGCGCGCAAGAGTGTACATGATGAAGCAACCTGGTGGGGCGCGCGAG 126
DB	
QY 121	CTGGAGAGTTGCGCCCTGAACCCCGCGCTGCTGGAGAGCGAGGCTGCAAGCAGATC 180
DB	
QY 127	CTGGAGAGTTGCGCCCTGAACCCCGCGCTGCTGGAGAGCGAGGCTGCAAGCAGATC 186
DB	
QY 181	ATCCGCCAGGTGCACCCCGCGCTGCAGACCGCGCAGGAGTCCGGACACCAAGAGGCC 240
DB	
QY 187	ATCCGCCAGGTGCACCCCGCGCTGCAGACCGCGCAGGAGTCCGGACACCAAGAGGCC 246
DB	
QY 241	ACCGTGGCCACCCCTGTACTGCGTGCACGAGAGATCCAGGTCCGGACACCAAGAGGCC 300
DB	
QY 247	ACCGTGGCCACCCCTGTACTGCGTGCACGAGAGATCCAGGTCCGGACACCAAGAGGCC 306
DB	
QY 301	CTGGACAAGATCGAGGAGGAGCAGACAAGTGCAGAGAGATCCAGCAGGCCCGAGGCC 360
DB	
QY 307	CTGGACAAGATCGAGGAGGAGCAGACAAGTGCAGAGAGATCCAGCAGGCCCGAGGCC 366
DB	
QY 361	GCCGACAAGGGCAAGGTGAGCCAGAGATACCCCATCCGTCAGAACCTGCAGGGCCAGATG 420
DB	
QY 367	GCCGACAAGGGCAAGGTGAGCCAGAGATACCCCATCCGTCAGAACCTGCAGGGCCAGATG 426
DB	
QY 421	GTGCACCGCCCATCAGCCCGCGCACCGCTGAACCGCTGGGTGAAGTTCGAGGAGAAG 480
DB	
QY 427	GTGCACCGCCCATCAGCCCGCGCACCGCTGAACCGCTGGGTGAAGTTCGAGGAGAAG 486
DB	
QY 481	GCCTTCAGCCCGCGAGGTGATCCCGATGTCACCGCGCTGAGCGAGCGGCCACCCCGCAG 540
DB	
QY 487	GCCTTCAGCCCGCGAGGTGATCCCGATGTCACCGCGCTGAGCGAGCGGCCACCCCGCAG 546
DB	
QY 541	GACCTGAACACGATGTTGACACCGCTGGCGCGCGCACCGCCCGCATGCAGATGTTGAAG 600
DB	
QY 547	GACCTGAACACGATGTTGACACCGCTGGCGCGCGCACCGCCCGCATGCAGATGTTGAAG 606
DB	
QY 601	GACACCATCAGCAGGAGCGCGCGAGTGGGACCGCTGCACCCCGTGCACCGCGCGCGCC 660
DB	
QY 607	GACACCATCAGCAGGAGCGCGCGAGTGGGACCGCTGCACCCCGTGCACCGCGCGCGCC 666
DB	
QY 661	ATCGCCCCCGCGCAGATGCGCGAGCGCGCGCGCACCATCGCGCGCACCCACCCAGCC 720
DB	
QY 667	ATCGCCCCCGCGCAGATGCGCGAGCGCGCGCGCACCATCGCGCGCACCCACCCAGCC 726
DB	
QY 721	CTGCAGGAGCAGATCGCCCTGGATGACCGACAACCCCGCATCCCGTGGCGCATCTAC 780
DB	
QY 727	CTGCAGGAGCAGATCGCCCTGGATGACCGACAACCCCGCATCCCGTGGCGCATCTAC 786
DB	
QY 781	AAGCGGIGGATCATCTCTGGCCCTGACCAAGATCGTGGGATGTACAGCCCGCTGAGCATC 840
DB	
QY 787	AAGCGGTGGATCATCTCTGGCCCTGACCAAGATCGTGGGATGTACAGCCCGCTGAGCATC 846
DB	

RESULT 10
US-10-190-435-20
; Sequence 20, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGERDE, Jan
; APPLICANT: BARNEFT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRUCHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190.435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 2742
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GagTatRevNef_C
US-10-190-435-20

Query Match 99.8%; Score 1475.8; DB 12; Length 2742;
Best Local Similarity 99.9%; Pred. No. 1.3e-291;
Matches 1477; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 241 ACCGTGGCCACCCCTGTAAGTGGTGCACGAGAGATCGAGGTTCGGGACACCAAGGAGGCC 300
DB
QY 247 ACCGTGGCCACCCCTGTAAGTGGTGCACGAGAGATCGAGGTTCGGGACACCAAGGAGGCC 306
DB
QY 301 CTGGACAAGATCGAGGAGGAGCAGACAAGTGCAGCAGAGAATCCAGCAGGCGGAGGCC 360
DB
QY 307 CTGGACAAGATCGAGGAGGAGCAGACAAGTGCAGCAGAGAATCCAGCAGGCGGAGGCC 366
DB
QY 361 GCCGACAAGGGCAAGGTGAGCCAGAACACTACCCCATGTCAGACACCTGCAGGCGCCAGATG 420
DB
QY 367 GCCGACAAGGGCAAGGTGAGCCAGAACACTACCCCATGTCAGACACCTGCAGGCGCCAGATG 426
DB
QY 421 GTGCACAGGCCATCAGGCGCCGACCCCTGAACCCCTGGGTGAAGGTGATCCAGGAGAG 480
DB
QY 427 GTGCACAGGCCATCAGGCGCCGACCCCTGAACCCCTGGGTGAAGGTGATCCAGGAGAG 486
DB
QY 481 GCCTTCAGCCCGAGGTGATCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCGCCAG 540
DB
QY 487 GCCTTCAGCCCGAGGTGATCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCGCCAG 546
DB
QY 541 GACCTGAACACGATCTTGAACACCCGTGGGCGGACACAGCGCCCATGTCAGATGCTGAAG 600
DB
QY 547 GACCTGAACACGATCTTGAACACCCGTGGGCGGACACAGCGCCCATGTCAGATGCTGAAG 606
DB
QY 601 GACACCATCAACGAGGAGCGCGGAGTGGGACCGGTGCACCCGCTGCAGCGGGCGCC 660
DB
QY 607 GACACCATCAACGAGGAGCGCGGAGTGGGACCGGTGCACCCGCTGCAGCGGGCGCC 666
DB
QY 661 ATGCGCCCGCGCCAGATGCGGAGGAGCCCGCGGACGACATCGCCGCGCACCGACCC 720
DB
QY 667 ATGCGCCCGCGCCAGATGCGGAGGAGCCCGCGGACGACATCGCCGCGCACCGACCC 726
DB
QY 721 CTGAGGAGCAGATCGCCTGGATGACACGACCAACCCCCCATCCCGGTGGGACATCTAG 780
DB
QY 727 CTGAGGAGCAGATCGCCTGGATGACACGACCAACCCCCCATCCCGGTGGGACATCTAG 786
DB
QY 781 AAGCGGTGGATCATCTCTGGCCTGAACAAGATCTGCGGATGTACAGCCCGCTGAGCATC 840
DB
QY 787 AAGCGGTGGATCATCTCTGGCCTGAACAAGATCTGCGGATGTACAGCCCGCTGAGCATC 846
DB
QY 841 CTGCACATCAAGCAGGCGCCCAAGAGCCCTTCGCGACTACTCTGGACCGCTTCTTCAAG 900
DB
QY 847 CTGCACATCAAGCAGGCGCCCAAGAGCCCTTCGCGACTACTCTGGACCGCTTCTTCAAG 906
DB
QY 901 ACCCTGCGCGCGAGCAGACACCCAGGAGGTGAAGAACTGGATGACCGACACCCCTGCTG 960
DB
QY 907 ACCCTGCGCGCGAGCAGACACCCAGGAGGTGAAGAACTGGATGACCGACACCCCTGCTG 966
DB
QY 961 GTGCAGAACGCCAACCCGACTGCAAGACCACTCTGCGCGCTCTCGGCCCGCGGCGAGC 1020
DB
QY 967 GTGCAGAACGCCAACCCGACTGCAAGACCACTCTGCGCGCTCTCGGCCCGCGGCGAGC 1026
DB
QY 1021 CTGGAGGAGATGATGACCGCCTGCGAGGGCGTGGCGCGCCCAAGCCACAAAGCGCGCTG 1080
DB
QY 1027 CTGGAGGAGATGATGACCGCCTGCGAGGGCGTGGCGCGCCCAAGCCACAAAGCGCGCTG 1086
DB
QY 1081 CTGGCGCAGCGATGAGCCAGCGCCAAACCCAGCGTGTATGATGAGAAAGAGCAACTTCAAG 1140
DB
QY 1087 CTGGCGCAGCGATGAGCCAGCGCCAAACCCAGCGTGTATGATGAGAAAGAGCAACTTCAAG 1146
DB
QY 1141 GGCCCGCGCGCATGTCAAAGTCTCAACTCGCGCAAGAGGGCGCCACATCGCCCGCAAG 1200
DB
QY 1147 GGCCCGCGCGCATGTCAAAGTCTCAACTCGCGCAAGAGGGCGCCACATCGCCCGCAAC 1206
DB
QY 1201 TGCCGCGCCCGCCGCAAGAAGGGTCTGGAAGTGGCGCAAGGAGGGGCCACCATGAG 1260
DB
QY 1207 TGCCGCGCCCGCCGCAAGAAGGGTCTGGAAGTGGCGCAAGGAGGGGCCACCATGAG 1266
DB
QY 1261 GACTGCACCGAGCGCGCAGGCCAACTTCTCTGGCAAGATCTGGGCCAGCCACAAAGGGCGC 1320
DB
QY 1267 GACTGCACCGAGCGCGCAGGCCAACTTCTCTGGCAAGATCTGGGCCAGCCACAAAGGGCGC 1326
DB

QY 1321 CCGGCAACTTCTCTGAGAGAGCCCGCGGAGCCACCGCCCGCCCGCGGAGAGCTTCGC 1380
DB
QY 1327 CCGGCAACTTCTCTGAGAGAGCCCGCGGAGCCACCGCCCGCCCGCGGAGAGCTTCGC 1386
DB
QY 1381 TCGAGGAGACACCCCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
DB
QY 1387 TCGAGGAGACACCCCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1446
DB
QY 1441 CTGAAGAGCCTGTTCTGGCAACGACCCCTGAGCCAGTAA 1479
DB
QY 1447 CTGAAGAGCCTGTTCTGGCAACGACCCCTGAGCCAGAA 1485
DB
RESULT 12
US-10-190-435-10
: Sequence 10, Application US/10190435
: Publication No. US20030143248A1
: GENERAL INFORMATION:
: APPLICANT: ZUR MEGEDE, Jan
: APPLICANT: BARNETT, Susan W.
: APPLICANT: LIAN, Ying
: APPLICANT: ENGELBRECHT, Susan
: APPLICANT: VAN RENSBURG, Estrelita J.
: TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
: TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES AND USES THEREOF
: FILE REFERENCE: PP18133.003 / 2302-18133
: CURRENT APPLICATION NUMBER: US/10/190,435
: CURRENT FILING DATE: 2002-12-30
: NUMBER OF SEQ ID NOS: 319
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 10
: LENGTH: 3930
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: GagComplPolmutAtt_C
US-10-190-435-10
Query Match 99.8%; Score 1475.8; DB 12; Length 3930;
Rest Local Similarity 99.9%; Pred. No. 1.2e-291;
Matches 1477; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGGCGCCCGCGGCGGAGATCTCTGCGCGGCGGAGAGCTGGAGCGCTGGGAGCGCATCCGC 60
DB
QY 7 ATGGGCGCCCGCGGCGGAGATCTCTGCGCGGCGGAGAGCTGGAGCGCTGGGAGCGCATCCGC 66
DB
QY 61 CTGGCGCCCGCGGCGGAGAGAGTGTACATGATGAAGACACCTTGTGTGGCGGAGCGCGAG 120
DB
QY 67 CTGGCGCCCGCGGCGGAGAGAGTGTACATGATGAAGACACCTTGTGTGGCGGAGCGCGAG 126
DB
QY 121 CTGGAGAGTTGCGCCTGAACCCCGGCGCTGTGGAGACACCGAGGCGCTGCAAGCAGATC 180
DB
QY 127 CTGGAGAGTTGCGCCTGAACCCCGGCGCTGTGGAGACACCGAGGCGCTGCAAGCAGATC 186
DB
QY 181 ATCCGCGAGCTGCACCCCGCCCTGCAGACACCGGCGAGGAGGCTGAAGAGCCCTGTCAAC 240
DB
QY 187 ATCCGCGAGCTGCACCCCGCCCTGCAGACACCGGCGAGGAGGCTGAAGAGCCCTGTCAAC 246
DB
QY 241 ACCGTGCGCACCCCTGTACTGCGTGCATGAGAGAGATCGAGGCTCCGGCCACACCAAGAGGCC 300
DB
QY 247 ACCGTGCGCACCCCTGTACTGCGTGCATGAGAGAGATCGAGGCTCCGGCCACACCAAGAGGCC 306
DB
QY 301 CTGGACAAGATCGAGGAGGAGCAGACAAGTCCAGCAGACAAGATCCAGAGGCGCGAGGCC 360
DB
QY 307 CTGGACAAGATCGAGGAGGAGCAGACAAGTCCAGCAGACAAGATCCAGAGGCGCGAGGCC 366
DB
QY 361 GCCGACAAGGGCAAGGTGAGCCAGAGACTACCCCATCTGTCAGACACCTGCAGGGCCAGATG 420
DB
QY 367 GCCGACAAGGGCAAGGTGAGCCAGAGACTACCCCATCTGTGCAGAACTCTGCAGGGCGAGATG 426
DB
QY 421 GTGCACAGGCCCATACACCCCGCGGAGCGCTGAACCGCTGGTGAAGGTGATCGAGGAGAG 480
DB
QY 427 GTGCACAGGCCCATACACCCCGCGGAGCGCTGAACCGCTGGTGAAGGTGATCGAGGAGAG 486
DB

```
QY 481 GCCTTCAGCCCGGAGGTGATCCCATGTTACCGCGCTGAGGAGGCGCCACCCCGCAG 540
Db 487 GCCTTCAGCCCGGAGGTGATCCCATGTTACCGCGCTGAGGAGGCGCCACCCCGCAG 546
QY 541 GACCTGAACACAGATGTTGAACACACCGTGGCGGGCCACAGGCGCCCATGAGAGTGAAG 600
Db 547 GACCTGAACACAGATGTTGAACACACCGTGGCGGGCCACAGGCGCCCATGAGAGTGAAG 606
QY 601 GACACCATCAACGAGGAGCGCGCGAGTGGGACCGCGTGCACCCCGTGCACCGCGCGCC 660
Db 607 GACACCATCAACGAGGAGCGCGCGAGTGGGACCGCGTGCACCCCGTGCACCGCGCGCC 666
QY 661 ATCGCCCGCGCCAGATGCGCGAGCGCGCGGAGCGGACATCGCGGACCGGACCGGACCG 720
Db 667 ATCGCCCGCGCCAGATGCGCGAGCGCGCGGAGCGGACATCGCGGACCGGACCGGACCG 726
QY 721 CTGAGGAGCAGATCGCGTGGATGACCAAGCAACCCCGCATCCCGTGGCGGACATCTAC 780
Db 727 CTGAGGAGCAGATCGCGTGGATGACCAAGCAACCCCGCATCCCGTGGCGGACATCTAC 786
QY 781 AAGCGGTGATCATCTGGGCGCTGAACAAAGATCGTGGGATGTACAGCCCGGTGAGCATC 840
Db 787 AAGCGGTGATCATCTGGGCGCTGAACAAAGATCGTGGGATGTACAGCCCGGTGAGCATC 846
QY 841 CTGGACATCAAGCAGGGCGCCCAAGGAGCGCTTCGGCGACTACGTGGACCGCTTCITCAAG 900
Db 847 CTGGACATCAAGCAGGGCGCCCAAGGAGCGCTTCGGCGACTACGTGGACCGCTTCITCAAG 906
QY 901 ACCCTGGCGCGGAGCAGACACCCAGGAGGTGAAGAACTGGATGACCGACACCCCTGCTG 960
Db 907 ACCCTGGCGCGGAGCAGACACCCAGGAGGTGAAGAACTGGATGACCGACACCCCTGCTG 966
QY 961 GTGCAGAACCGCCAAACCCCGACTGCAAGACCATCTCGCGGCTCTCGGCCCGCGCGGAGC 1020
Db 967 GTGCAGAACCGCCAAACCCCGACTGCAAGACCATCTCGCGGCTCTCGGCCCGCGCGGAGC 1026
QY 1021 CTGGAGGAGATGATGACCGCTGCCAGGGCGTGGCGGCCCGCCAGCCACAGGCGCGTG 1080
Db 1027 CTGGAGGAGATGATGACCGCTGCCAGGGCGTGGCGGCCCGCCAGCCACAGGCGCGTG 1086
QY 1081 CTGGCGGAGCGGATGAGCCAGGCGCAACACCGAGCTGATGATGAGAGAGCAACTTCAAG 1140
Db 1087 CTGGCGGAGCGGATGAGCCAGGCGCAACACCGAGCTGATGATGAGAGAGCAACTTCAAG 1146
QY 1141 GSCCGCGCGCGCATCGTCAAGTGTCTTCACTGCGGCAAGGAGGCGCCACATCGCCGCAAG 1200
Db 1147 GSCCGCGCGCGCATCGTCAAGTGTCTTCACTGCGGCAAGGAGGCGCCACATCGCCGCAAG 1206
QY 1201 TGCCGCGCGCGCGCATCGTCAAGTGTCTTCACTGCGGCAAGGAGGCGCCACATCGAGTGAAG 1260
Db 1207 TGCCGCGCGCGCGCATCGTCAAGTGTCTTCACTGCGGCAAGGAGGCGCCACATCGAGTGAAG 1266
QY 1261 GACTGCACCGAGCGCGCAAGGCGTCTGGAAGTGGGCAAGGAGGCGCCACATCGAGTGAAG 1320
Db 1267 GACTGCACCGAGCGCGCAAGGCGTCTGGAAGTGGGCAAGGAGGCGCCACATCGAGTGAAG 1326
QY 1321 CCCGGCAACTTCTTGCAAGTCCCGGCGCCAGCCACCGCGCGCGCGCGCGCGCGCGCGCG 1380
Db 1327 CCCGGCAACTTCTTGCAAGTCCCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1386
QY 1381 TTCGAGGAGACCCACCCCGCGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
Db 1387 TTCGAGGAGACCCACCCCGCGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1446
QY 1441 CTGAAGAGCGCTGTTGCGGCAACGACCCCGCTGAGCGCACTAA 1479
Db 1447 CTGAAGAGCGCTGTTGCGGCAACGACCCCGCTGAGCGCACTAA 1485
```

RESULT 13
US-10-190-435-11
; Sequence 11, Application US/10190435

```
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jac
; APPLICANT: HARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 3930
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GagComplPolmutina_C
US-10-190-435-11
```

```
Query Match 99.8%; Score 1475.8; DB 12; Length 3930;
Best Local Similarity 99.9%; Pred. No. 1.2e-291;
Matches 1477; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGGGCGCGCGCCAGCATCTCTGCGCGGGCGGCAAGCTTGACGCGCTGGGAGCGCATCCGC 60
Db 7 ATGGGGCGCGCGCCAGCATCTCTGCGCGGGCGGCAAGCTTGACGCGCTGGGAGCGCATCCGC 66
QY 61 CTGCGCGCGCGCGGCAAGAGTGTCTACATGATGAAGCACCCTGGTGTGGCCAGCGCGAG 120
Db 67 CTGCGCGCGCGCGGCAAGAGTGTCTACATGATGAAGCACCCTGGTGTGGCCAGCGCGAG 126
QY 121 CTGGAGAAGTTCCGCTTGAACCCCGCGCTTGGAGACCAGGAGGCTGCAAGCAGATC 180
Db 127 CTGGAGAAGTTCCGCTTGAACCCCGCGCTTGGAGACCAGGAGGCTGCAAGCAGATC 186
QY 181 ATCCGCCAGCTGCACCCCGCGCTGCAGAGCGGCGGAGAGAGCTGAAGAGCGCTTCAAC 240
Db 187 ATCCGCCAGCTGCACCCCGCGCTGCAGAGCGGCGGAGAGAGCTGAAGAGCGCTTCAAC 246
QY 241 ACCGTGGCCACCTGTACTTCTGCTCCAGGAGAGATCGAGGTCCCGGACACCAAGGAGCC 300
Db 247 ACCGTGGCCACCTGTACTTCTGCTCCAGGAGAGATCGAGGTCCCGGACACCAAGGAGCC 306
QY 301 CTGGACAAGATCGAGGAGGAGGAGAGAGAGTGCAGAGAGAGATCCAGAGGCGCGAGGCC 360
Db 307 CTGGACAAGATCGAGGAGGAGGAGAGAGAGTGCAGAGAGAGATCCAGAGGCGCGAGGCC 366
QY 361 GCGGACAAGGCAAGGTGAGCGAGAGAGTGCAGAGAGTGCAGAGAGTGCAGAGGCGCGAGT 420
Db 367 GCGGACAAGGCAAGGTGAGCGAGAGAGTGCAGAGAGTGCAGAGAGTGCAGAGGCGCGAGT 426
QY 421 GTGCACCGAGCGCATCAGCCCGCGCACCTTGAACGCTTGGTGAAGGTGATCGAGGAGAG 480
Db 427 GTGCACCGAGCGCATCAGCCCGCGCACCTTGAACGCTTGGTGAAGGTGATCGAGGAGAG 486
QY 481 GCGTTTCAGCCCGGAGGTGATCCCGCATGTTTACCGCGCGCTGAGCGAGGCGCGACCCCGCAG 540
Db 487 GCGTTTCAGCCCGGAGGTGATCCCGCATGTTTACCGCGCGCTGAGCGAGGCGCGACCCCGCAG 546
QY 541 GACCTGAACACAGATGTTGAACACCGTGGCGGGCGCCACAGGCGCCCATGAGAGTGAAG 600
Db 547 GACCTGAACACAGATGTTGAACACCGTGGCGGGCGCCACAGGCGCCCATGAGAGTGAAG 606
QY 601 GACACCAICACGAGGAGCGCGGAGTGGGAGCGCGTGCACCCCGTGCACCGCGCGCGCC 660
Db 607 GACACCAICACGAGGAGCGCGGAGTGGGAGCGCGTGCACCCCGTGCACCGCGCGCGCC 666
QY 661 ATCGCCCGCGCGCAGATGCGCGGAGCGCGCGGCGGAGAGATCGCGGCGGAGAGAGAGCACC 720
Db 667 ATCGCCCGCGCGCAGATGCGCGGAGCGCGCGGCGGAGAGATCGCGGCGGAGAGAGAGCACC 726
```

```
QY 721 CTGCAGGAGCAGATCGCTGGATGACCAACACCCGCCCATCCCTGGCGGACATCTAC 780
      |||
Db 727 CTGCAGGAGCAGATCGCTGGATGACCAACACCCGCCCATCCCTGGCGGACATCTAC 786
      |||
QY 781 AAGCGGTGGATCATCTCTGGCCCTGAACAAGATCTGCGGATGTACAGCCCGCTGAGCATC 840
      |||
Db 787 AAGCGGTGGATCATCTCTGGCCCTGAACAAGATCTGCGGATGTACAGCCCGCTGAGCATC 846
      |||
QY 841 CTGGACATCAAGCAGGCGCCCAAGGAGCCCTTCGCGGACTAGCTGGACCGCTTCCTCAAG 900
      |||
Db 847 CTGGACATCAAGCAGGCGCCCAAGGAGCCCTTCGCGGACTAGCTGGACCGCTTCCTCAAG 906
      |||
QY 901 ACCCTGCGCGCGGAGCAGACCCAGGAGCTGGAAGAACTGGATGATGACCGACCCCTGCTG 960
      |||
Db 907 ACCCTGCGCGCGGAGCAGACCCAGGAGCTGGAAGAACTGGATGATGACCGACCCCTGCTG 966
      |||
QY 961 GTGCAGAAAGCCAAACCCCGGACTGCAAGAGCCATCTCTGCGGGCTCTGGCCCGCGCGCAGC 1020
      |||
Db 967 GTGCAGAAAGCCAAACCCCGGACTGCAAGAGCCATCTCTGCGGGCTCTGGCCCGCGCGCAGC 1026
      |||
QY 1021 CTGCAGGAGATGATGACCGCCCTGCCAGGGCGTGGCGGCGCCCGCCAGCCACAAAGCCCGGTC 1080
      |||
Db 1027 CTGCAGGAGATGATGACCGCCCTGCCAGGGCGTGGCGGCGCCCGCCAGCCACAAAGCCCGGTC 1086
      |||
QY 1081 CTGGCGGAGCGGATGAGCCAGCGCCCAACACAGCGGTGATGATGATGATGATGATGATGATGATG 1140
      |||
Db 1087 CTGGCGGAGCGGATGAGCCAGCGCCCAACACAGCGGTGATGATGATGATGATGATGATGATGATG 1146
      |||
QY 1141 GGCCTGCGCGGCGCATCTCAAGTGTCTCAACTGGCGCAAGGAGGCGCCACATGCGCGCGCAAC 1200
      |||
Db 1147 GGCCTGCGCGGCGCATCTCAAGTGTCTCAACTGGCGCAAGGAGGCGCCACATGCGCGCGCAAC 1206
      |||
QY 1201 TGCCTGCGCGCGCGCAAGAGGGCTGCTGGAAGTGGCGCAAGGAGGCGCCACAGATGAAG 1260
      |||
Db 1207 TGCCTGCGCGCGCGCAAGAGGGCTGCTGGAAGTGGCGCAAGGAGGCGCCACAGATGAAG 1266
      |||
QY 1261 GACTGCACCGGAGCGCGCAGGCGCAACTCTCTGGCAAGATCTGGCCAGCGCCACAAAGGCGCGC 1320
      |||
Db 1267 GACTGCACCGGAGCGCGCAGGCGCAACTCTCTGGCAAGATCTGGCCAGCGCCACAAAGGCGCGC 1326
      |||
QY 1321 CCGGCAACTTCTTCAGAGCGCGCGCCCGGAGCCCGCCCGCCCGCGCGGAGCTTCGCGC 1380
      |||
Db 1327 CCGGCAACTTCTTCAGAGCGCGCGCGCGGAGCCCGCGCGCGCGCGCGGAGCTTCGCGC 1386
      |||
QY 1381 TTCGAGGAGACCGCGCGCGCGCGCAAGAGGAGCAAGGAGCGCGGAGACCGCTGACCCAGC 1440
      |||
Db 1387 TTCGAGGAGACCGCGCGCGCGCGCAAGAGGAGCAAGGAGCGCGGAGACCGCTGACCCAGC 1446
      |||
QY 1441 CTGAAGAGCGCTGTTGGGCAACGACCCCTGAGCCAGTAA 1479
      |||
Db 1447 CTGAAGAGCGCTGTTGGGCAACGACCCCTGAGCCAGTAA 1485
      |||
```

```
RESULT 14
US-10-190-435-12
; Sequence 12, Application: US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 5145
```

```
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: GagComplPolmutInatRevNef_C
US-10-190-435-12

Query Match          99.8%; Score 1475.8; DB 12; Length 5145;
Best Local Similarity 99.9%; Pred. No. 1.2e-291;
Matches 1477; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGGCGCCCGCGCGCAGCATCTCTGGCGCGCGGCGGCAAGCTGGACGCTGGAGCGCATCCGC 60
      |||
Db 7 ATGGGCGCCCGCGCGCAGCATCTCTGGCGCGCGGCGGCAAGCTGGACGCTGGAGCGCATCCGC 66
      |||
QY 61 CTGGCGCCCGCGCGGCAAGAGTGTATACATGATGAAGCACTGTGGTGTGGCGCAGCCGCGAG 120
      |||
Db 67 CTGGCGCCCGCGCGGCAAGAGTGTATACATGATGAAGCACTGTGGTGTGGCGCAGCCGCGAG 126
      |||
QY 121 CTGGAGAAAGTTGCGCCTGAACCCCGCGCCTGCTGGAGACCAAGCGAGGGCTGCAAGCAGATC 180
      |||
Db 127 CTGGAGAAAGTTGCGCCTGAACCCCGCGCCTGCTGGAGACCAAGCGAGGGCTGCAAGCAGATC 186
      |||
QY 181 ATCCGCCAGCTGCACCCCGCGCCTGCAGACCCGCGGAGGAGGAGTGAAGAGCCCTGTTCAAC 240
      |||
Db 187 ATCCGCCAGCTGCACCCCGCGCCTGCAGACCCGCGGAGGAGGAGTGAAGAGCCCTGTTCAAC 246
      |||
QY 241 ACCGTGGCCACCCCTGTACTGCGTGCACGAGAACTGAGGTTCGCGGACACCAAGGAGGCC 300
      |||
Db 247 ACCGTGGCCACCCCTGTACTGCGTGCACGAGAACTGAGGTTCGCGGACACCAAGGAGGCC 306
      |||
QY 301 CTGRCAGAAGATCGAGGAGGAGCAGAACAAAGTGCAGCAGCAGAAATCCAGCAGGCCGAGGCC 360
      |||
Db 307 CTGRCAGAAGATCGAGGAGGAGCAGAACAAAGTGCAGCAGCAGAAATCCAGCAGGCCGAGGCC 366
      |||
QY 361 GCCGACAAGGGCAGGGTGAAGCAGAACTACCCCATCTGTCAGAACTTGCAGGSCCAGATG 420
      |||
Db 367 GCCGACAAGGGCAGGGTGAAGCAGAACTACCCCATCTGTCAGAACTTGCAGGSCCAGATG 426
      |||
QY 421 GTGCACCAAGGCGCATCAGCCCGCGCAGCCTGAACGCTGGTGAAGGTGATCGAGGAGGAAG 480
      |||
Db 427 GTGCACCAAGGCGCATCAGCCCGCGCAGCCTGAACGCTGGTGAAGGTGATCGAGGAGGAAG 486
      |||
QY 481 GCCTTCAGCCCGCGAGGTGATCCCATGTTCACCGCGCTGAGCGAGGCGCGCCACCCCGCAG 540
      |||
Db 487 GCCTTCAGCCCGCGAGGTGATCCCATGTTCACCGCGCTGAGCGAGGCGCGCCACCCCGCAG 546
      |||
QY 541 GACCTGAACACAGATGTGAACACCGTGGCGGCGCAACAGCGCGCATGCAGATGCTGAAG 600
      |||
Db 547 GACCTGAACACAGATGTGAACACCGTGGCGGCGCAACAGCGCGCATGCAGATGCTGAAG 606
      |||
QY 601 GACACCATCAACGAGGAGCGCGCGAGTGGGAGCGGTGCACCGCGTGCACCGCGCGCCCGC 660
      |||
Db 607 GACACCATCAACGAGGAGCGCGCGAGTGGGAGCGGTGCACCGCGTGCACCGCGCGCCCGC 666
      |||
QY 661 ATCGCCCGCGCGCGAGATGCGCGAGCCCGCGCGGAGCGCAATCGCGCGGACACCGACACC 720
      |||
Db 667 ATCGCCCGCGCGCGAGATGCGCGAGCCCGCGCGGAGCGCAATCGCGCGGAGCGCGCGCC 726
      |||
QY 721 CTGCAAGAGCAGATCGCTGGATGACCAAGCAACCGCCCGCCCATCCCGCTGGCGGACATCTAC 780
      |||
Db 727 CTGCAAGAGCAGATCGCTGGATGACCAAGCAACCGCCCGCCCATCCCGCTGGCGGACATCTAC 786
      |||
QY 781 AAGCGGTGGATCATCTCTGGCGCTGAACAAGATCTGCGGATGTACAGCCCGCTGAGCATC 840
      |||
Db 787 AAGCGGTGGATCATCTCTGGCGCTGAACAAGATCTGCGGATGTACAGCCCGCTGAGCATC 846
      |||
QY 841 CTGGACATCAAGCAGGCGCCCAAGGAGCCCTTCGCGGACTAGCTGGACCGCTTCTTCAAG 900
      |||
Db 847 CTGGACATCAAGCAGGCGCCCAAGGAGCCCTTCGCGGACTAGCTGGACCGCTTCTTCAAG 906
      |||
QY 901 ACCCTGCGCGCGGAGCAGACCCAGGAGGTGAAGAACTGGATGATGACCGACCCCTGCTG 960
      |||
```

Db 907 ACCCTGGCGCGGAGAGAGACCCAGGAGGTGAAGAACTGGATGACCGAGACCGTGGCTG 998
QY 961 GTGCAGAACGCCAACCCCGACTGCAAGAGACCATCTGCGGCTCTGCGCCCGCGCGCCAGC 1020
Db 967 GTGCAGAACGCCAACCCCGACTGCAAGAGACCATCTGCGGCTCTGCGCCCGCGCGCCAGC 1026
QY 1021 CTGAGGAGATGATGACCGCTTGCAGGGCGGTGGGCGGCCCCAGGCCACAAAGGCGCGGTC 1080
Db 1027 CTGAGGAGATGATGACCGCTTGCAGGGCGGTGGGCGGCCCCAGGCCACAAAGGCGCGGTC 1086
QY 1081 CTGCGCGAGGCGATGAGCCAGGCCAACACACCGGTGATGATGAGAGAGAGACACTTCAAG 1140
Db 1087 CTGCGCGAGGCGATGAGCCAGGCCAACACACCGGTGATGATGAGAGAGAGACACTTCAAG 1146
QY 1141 GGCCCCCGCGCATCGTCAAGTGTTCAACTGCGGCAAGGAGGGCCACATCGGCCCGCAAC 1200
Db 1147 GGCCCCCGCGCATCGTCAAGTGTTCAACTGCGGCAAGGAGGGCCACATCGGCCCGCAAC 1206
QY 1201 TGCCCGCCCCCGCAAGAGGGCTGCTGGAAGTGGCGGCAAGGAGGGCCACCAATGAAG 1260
Db 1207 TGCCCGCCCCCGCAAGAGGGCTGCTGGAAGTGGCGGCAAGGAGGGCCACCAATGAAG 1266
QY 1261 GACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGGCCAGCCACACAGGGCGCG 1320
Db 1267 GACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGGCCAGCCACACAGGGCGCG 1326
QY 1321 CCGGCAACTTCCTGCAGAGCGCGCGCCAGCCACCGCCCGCCCGCGCGAGAGCTTCGGC 1380
Db 1327 CCGGCAACTTCCTGCAGAGCGCGCGCCAGCCACCGCCCGCCCGCGCGAGAGCTTCGGC 1386
QY 1381 TTCGAGGAGACCAACCCCGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Db 1387 TTCGAGGAGACCAACCCCGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1446
QY 1441 CTGAAGAGCCTGTTCCGGCAACGACCCCTTGAGCCAGTAA 1479
Db 1447 CTGAAGAGCCTGTTCCGGCAACGACCCCTTGAGCCAGTAA 1485

RESULT 15

US-10-190-435-59
; Sequence 59, Application: US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEYER, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Esmerita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PPI8133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59
; LENGTH: 4713
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: tatRevNeiGagProteinArtmut C
US-10-190-435-59

Query Match 99.7%; Score 1474.8; DB 12; Length 4713;
Best Local Similarity 99.9%; Pred. No. 2e-291;
Matches 1476; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 TGGCGCGCGCGCGCGAGCATCTGCGGCGCGGCAAGCTGGACGCTGGGAGCGCATCGCC 61
Db 1262 TTGGCGCGCGCGCGCGAGCATCTGCGGCGCGGCAAGCTGGACGCTGGGAGCGCATCGCC 1321

QY 52 TGGCGCGCGCGCGCGAGCATCTGCTAATGATGAAGCACTGCTGTGGGCGCAAGCGCGAGC 121
Db 1322 TGGCGCGCGCGCGCGAGCATCTGCTAATGATGAAGCACTGCTGTGGGCGCAAGCGCGAGC 1381
QY 122 TGGAGAGTTTGGCCCTGCAACCCCGGCTGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAG 181
Db 1382 TGGAGAGTTTGGCCCTGCAACCCCGGCTGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1441
QY 182 TCGCCAGCTGCACCCCGCTTGCAGACCGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 241
Db 1442 TCGCCAGCTGCACCCCGCTTGCAGACCGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1501
QY 242 CCGTGGCCACCTGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 301
Db 1502 CCGTGGCCACCTGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1561
QY 302 TGGACAGATCGAGGAG 361
Db 1562 TGGACAGATCGAGGAG 1621
QY 362 CCGACAGAGGCAAGGTGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 421
Db 1622 CCGACAGAGGCAAGGTGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1681
QY 422 TGCACAGAGGCAAGGTGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 481
Db 1682 TGCACAGAGGCAAGGTGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1741
QY 482 CCTCAGCGCGGAGGTGATCCCGATGATCCCGATGATCCCGATGATCCCGATGATCCCGAT 541
Db 1742 CCTCAGCGCGGAGGTGATCCCGATGATCCCGATGATCCCGATGATCCCGATGATCCCGAT 1801
QY 542 ACCTGAACAGCATGTTGAACACCGTGGCGGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAG 601
Db 1802 ACCTGAACAGCATGTTGAACACCGTGGCGGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAG 1861
QY 602 ACACCATCAACAGGAGGCGCGCGAGTGGAGCGCGTGGAGCGCGTGGAGCGCGTGGAGCG 661
Db 1862 ACACCATCAACAGGAGGCGCGCGAGTGGAGCGCGTGGAGCGCGTGGAGCGCGTGGAGCG 1921
QY 662 TCGCCCCCGCGCGAGATGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 721
Db 1922 TCGCCCCCGCGCGAGATGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1981
QY 722 TGCAGGAGAGATCGCTTGGATGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 781
Db 1982 TGCAGGAGAGATCGCTTGGATGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2041
QY 782 AGCGTGGATCATCTTGGGCTTGAACAGAGATGCTGGGAGAGAGAGAGAGAGAGAGAGAGAG 841
Db 2042 AGCGTGGATCATCTTGGGCTTGAACAGAGATGCTGGGAGAGAGAGAGAGAGAGAGAGAGAG 2101
QY 842 TGGACATCAAGCAGGCG 901
Db 2102 TGGACATCAAGCAGGCG 2161
QY 902 CCTTGGCGCGCGAGCAG 961
Db 2162 CCTTGGCGCGCGAGCAG 2221
QY 962 TGCAG 1021
Db 2222 TGCAG 2281
QY 1022 TGGAGGAGATGATGACCGCTGCCAGGGCGTGGCGGCGCGCGCGCGCGCGCGCGCGCGCG 1081
Db 2282 TGGAGGAGATGATGACCGCTGCCAGGGCGTGGCGGCGCGCGCGCGCGCGCGCGCGCGCG 2341
QY 1082 TGGCGCGAGGAGATGAGCCAGGCGCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1141
Db 2342 TGGCGCGAGGAGATGAGCCAGGCGCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2401
QY 1142 GCGCGCGCGCGCGATCGTCAAGTGTCTTCAACTGCGGCAAGGAGGCGCGCGCGCGCGCGCA 1201

```

Db      2402  GCGGCGGCGCATGTCAGTGGTCAACTGCGGCAAGGAGGGCCACATCGCGCGGAACT 2461
QY      1202  GCGGCGGCGCGCGGCAAGAGGGGCTGCTGAAGTGGGGCAAGGAGGGCCACAGATGAAGG 1261
Db      2462  GCGGCGGCGCGCGGCAAGAGGGGCTGCTGAAGTGGGGCAAGGAGGGCCACAGATGAAGG 2521
QY      1262  ACTGCACGAGCGCCAGGCGCAACTTCCTGGGCAAGATGTGGCCCGAGCCCAAGGGCGCGG 1321
Db      2522  ACTGCACGAGCGCCAGGCGCAACTTCCTGGGCAAGATGTGGCCCGAGCCCAAGGGCGCGG 2581
QY      1322  CCGGCAACTTCCTGCAGAGCGCGCGCGGAGCGCCCGCGCGCGCGCGCGGAGCTTCGGCT 1381
Db      2582  CCGGCAACTTCCTGCAGAGCGCGCGCGGAGCGCCCGCGCGCGCGCGCGGAGCTTCGGCT 2641
QY      1382  TCGAGGAGACACCGCGCGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1441
Db      2642  TCGAGGAGACACCGCGCGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2701
QY      1442  TGAAGAGCGCTGTCGGCAACGACCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1501
Db      2702  TGAAGAGCGCTGTCGGCAACGACCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2761
```

Search completed: September 23, 2003, 02:00:04
Job time : 375.76 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Computer, Ltd.

QM nucleic - nucleic search, using sw model

Run on: September 22, 2003, 17:33:23 : Search time 95.5311 Seconds
(without alignments)
6833.434 Million cell updates/sec

Title: US-09-475-704A-3
Perfect score: 1479
Sequence: 1 atggggcgcgcgcgcagcat.....acgaccccccgtgacccagtaa 1479

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 6
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_5/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	947.6	64.1	4307	4	US-09-552-950-2
2	939.8	63.5	9772	4	US-09-552-950-5
3	738.2	49.9	1476	4	US-09-184-418C-100
4	738.2	49.9	8959	4	US-09-184-418C-11
5	729.8	49.3	1485	4	US-09-184-418C-73
6	729.8	49.3	9010	4	US-09-184-418C-8
7	710.8	48.1	1486	4	US-09-184-418C-38
8	710.8	48.1	8992	4	US-09-184-418C-4
9	701.2	47.4	1496	4	US-09-184-418C-82
10	701.2	47.4	8972	4	US-09-184-418C-9
11	659.8	44.6	1503	4	US-09-184-418C-29
12	659.8	44.6	8953	4	US-09-184-418C-3
13	658.8	44.5	1479	4	US-09-184-418C-12
14	658.8	44.5	8968	4	US-09-184-418C-1
15	656.6	44.4	4307	4	US-09-184-418C-48
16	656.6	44.4	8966	4	US-09-184-418C-5
17	652.6	44.1	1494	4	US-09-184-418C-47
18	650.2	44.0	4307	4	US-09-552-950-1
19	650.2	44.0	9719	4	US-09-700-304-1
20	647	43.7	3807	2	US-08-417-210A-78
21	645.4	43.6	7399	2	US-08-418-848A-9
22	645.4	43.6	9709	2	US-08-188-583-5
23	645.4	43.6	9709	3	US-08-388-353-1
24	645.4	43.6	9709	3	US-08-488-551B-1
25	645.4	43.6	9709	4	US-09-309-572-15
26	645.4	43.6	12479	4	US-09-318-138-13
27	645.4	43.6	12494	3	US-08-935-312-13

28	645.4	43.6	12494	3	US-08-848-760B-33	Sequence 33, Appl
29	645.4	43.6	15581	3	US-08-646-538-35	Sequence 35, Appl
30	645.4	43.6	15581	3	US-09-503-222-35	Sequence 35, Appl
31	642.2	43.4	1503	4	US-09-184-418C-64	Sequence 64, Appl
32	642.2	43.4	9050	4	US-09-184-418C-7	Sequence 7, Appl
33	641.4	43.4	8987	4	US-09-184-418C-2	Sequence 2, Appl
34	641	43.3	9737	2	US-08-944-449-7	Sequence 7, Appl
35	641	43.3	9737	4	US-09-353-362-7	Sequence 7, Appl
36	637	43.1	1491	4	US-09-184-418C-91	Sequence 91, Appl
37	637	43.1	9060	4	US-09-184-418C-10	Sequence 10, Appl
38	633.2	42.7	5362	3	US-08-463-210-5	Sequence 5, Appl
39	628	42.5	8932	3	US-09-124-900-1	Sequence 1, Appl
40	628	42.5	8933	3	US-08-463-210-4	Sequence 4, Appl
41	628	42.5	8933	3	US-09-620-958A-3	Sequence 3, Appl
42	628	42.5	8933	3	US-09-620-958A-4	Sequence 4, Appl
43	628	42.5	8933	3	US-09-620-958A-9	Sequence 9, Appl
44	624	42.2	9746	1	US-08-022-835-3	Sequence 3, Appl
45	624	42.2	9746	1	US-08-388-809-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-09-552-950-2
: Sequence 2, Application: US/09552950
: Patent No. 6541248
: GENERAL INFORMATION:
: APPLICANT: Oxford Biomedica (UK) Limited
: TITLE OF INVENTION: Anti-Viral Vectors
: FID REFERENCE: 674524-2004
: CURRENT APPLICATION NUMBER: US/09/552,950
: CURRENT FILING DATE: 2000-04-20
: NUMBER OF SEQ ID NOS: 22
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 2
: LENGTH: 4307
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence:gagpol - codon
: OTHER INFORMATION: optimised gagpol sequence
US-09-552-950-2

Query Match	64.1%;	Score 947.6;	DB 4;	Length 4307;
Best Local Similarity	80.7%;	Pred. No. 9.2e-135;		
Matches 1135;	Conservative 0;	Mismatches 259;	Indels 12;	Gaps 2;
QY	1	ATGGGGCCCGCGCGCCAGCATCTCGCGGGCGGCAAGCTGGAGCGCCTGGGAGCGCATCCGC	60	
DB	1	ATGGGGCCCGCGCGCCAGCTCTGTGCGGGCGGAGCTGGACCGTGGGAGAGATCCGC	60	
QY	51	CTGGCGCCCGCGCGGCAAGAGTGTACATGATGAAGCACCTGGTGGGGCCAGCCGGAG	120	
DB	51	CTGGCGCCCGCGCGGCAAGAGTGTACATGATGAAGCACCTGTGTGGGGCCAGCCGGAA	120	
QY	121	CTGGAGAGTTTCGCCCTGAACCCCGGCTGCTGGAGACCCAGCGAGGGCTGCAAGCAGATC	180	
DB	121	CTGGAGAGTTTCGCCCTGAACCCCGGCTGCTGGAGACCCAGCGAGGGCTGCGCCAGATC	180	
QY	181	ATCCGCCAGCTGCACCCCGGCTGCGAGACCCGGCAGGAGGAGCTGAGAGCCCTGTTCAC	240	
DB	181	ATCCGCCAGCTGCACCCCGGCTGCGAGACCCGGCAGGAGGAGCTGCGAGCCCTGTTCAC	240	
QY	241	ACCGTGGCCACCTGTACTGGTGCACGAGAGATGAGGTCCGCGACACCAAGAGGCC	300	
DB	241	ACCGTGGCCACCTGTACTGGTGCACGAGAGATGAGGTCCGCGACACCAAGAGGCC	300	
QY	301	CTGGACAGATCGAGGAGGAGCAACAAGTCCAGCAGAGAGATCCAGCAGGCCAGGCC	360	
DB	301	CTGGATAAATCGAGGAGGAGCAACAAGTCCAGCAGAGAGATCCAGCAGGCCAGGCC	360	
QY	361	GCC-----GACAAAGGCAAGGTGAGCCAGAACTA:CCCCATCGTGCAGAACTTGCAGGGC	414	

```

Db      361 GACACCGGACACAGCAACAGGTCAGCCAGAACTACCCCACTCGTCAGAAACATCCAGGGC 429
Qy      415 CAGATGGTGCACACGACCATCAGCCCGCCGACCCCTGAACGCTGGGTGAAGGTGATCGAG 474
Db      421 CAGATGGTGCACACGACCATCTCCCGCCGACCGCTGAACGCTGGGTGAAGGTGATGGA 480
Qy      475 GAGAAGGCTTCAGCCCGGAGGTGATCCCATGTTTCAACGCGCCCTGAACGAGGGCGCCACC 534
Db      481 GAGAAGGCTTTAGCCCGGAGGTGATACCCATGTTTTCAGCCCTGTTCAGAGGGAGCCACC 540
Qy      535 CCCCAGGACCTGAACACGATGTTGAACACCGTGGCGGCCACCCAGCCCGCCATOCACATG 594
Db      541 CCCCAGATCTGAACACCATGCTCAACACAGTGGGGGACACCCAGCCCGCCATGACATG 600
Qy      595 CTGAAGGACACCATCAACGAGGAGGCGCGCGAGTGGGACCGCGTGCACCCCGTGCACGCC 654
Db      601 CTGAAGGAGACCATCAATGAGGAGGCTGCCGATGGGATCGTGTGCATCCGTCACCGCA 650
Qy      655 GGCCCCATGCCCGCCAGATGCCGAGCCCGCGGAGCCCGCGGACGAGATCGCCGGACACC 714
Db      661 GGCCCCATGCCCGCCAGATGCCGTGAGCCACCGCGGCTCAGACATCGCCGGACGACT 720
Qy      715 AGCACCTGTCAGGAGGACAGATCGCTGGGCTGAACAGATCGTGGGATGTACAGCCCGGTG 774
Db      721 AGTACCTTCAGGAACAGATCGCTGGATGACCAACACCCACCATCCCGTGGAGAA 780
Qy      775 AICTAAGCGGTGGATCATCTGGGCTGAACAGATCGTGGGATGTACAGCCCGGTG 834
Db      781 ATCTAAGACGCTGGATCATCTGGGCTGAACAGATCGTGGGATGTATAGCCCIACC 840
Qy      835 AGCATCTGGACATCAAGCAGGCGGCCCAAGGAGCCCTTCCGGACTACGTGGACCGCTC 894
Db      841 AGCATCTGGACATCCGCCAAGCGCCGAAGAACCCCTTTCGGGACTACGTGGACCGGTC 900
Qy      895 TTCAAGACCCCTGGCGCGGAGGACAGCAGCAGCCAGGAGGTGAAGAACTGCATGACGACATC 954
Db      901 TACAAGACCCCTGGCGCGGAGGAGGCTAGCCAGGAGGTGAAGAACTGATGACGCAACCC 950
Qy      955 CTGCTGGTGCAGAACCGCAACCCCGACTGCAAGACCATCTCTGGCGCTCTCGGCCCGCC 1014
Db      961 CTGCTGGTGCAGAACCGCAACCCCGACTGCAAGACCATCTCTGAAGCCCTTGGGCCAGCG 1020
Qy      1015 GCCAGCTTGAGGAGATGATGACCCGCTGCCAGGGCGTGGCGGCCCGCCCAAGGAGCC 1074
Db      1021 GCTACCTTAGAGGAATGATGACCCGCTGTCAAGGAGTGGCGGAGCCCGCCCAAGGCA 1080
Qy      1075 CGCGTGTGGCCGAGGAGATGAGCCAGGCCAACCA - - - - -CCAGGTGATGATCGAGAG 1128
Db      1081 CGCGTGTGTGCTGAGGCCATGAGCCAGGTGACCAACTCCGCTACCATCATGATGACGCGC 1140
Qy      1128 AGCAACTTCAAGGGCGCCCGCGGCTGCTCAAGTCTTCAACTGCGGCAAGGAGGCGCAC 1188
Db      1141 GGCAACTTTCGGAACCAACGCAAGATGCTCAAGTCTTCAACTGCTGGCAAGAAAGCGCAC 1200
Qy      1188 ATCCCGCGCAACTGCGCGCGCCCGCCGCAAGAGGGTGTGGAAGTTCGGCAAGGAGGCG 1248
Db      1201 ACAGCCCGCAACTGCGAGGGCGCCCTAGGAAAGGGCTGCTGGAAATGCGGCAAGGAGCG 1260
Qy      1249 CACCAATGAAGGACTGCAACCGCAGGCGCCAGGCAACTCTCTGGCAAGATCTGGCGCAGC 1308
Db      1261 CACCAATGAAGACTGTACGGAGAGACAGGCTAATTTTATAGGAAGATCTGGCCCTTC 1320
Qy      1309 CACAAGGGCGCCCGCCGCAACTTCTGCAAGAGCCCGCCGAGCCCGCCCGCCCGCCG 1368
Db      1321 TACAAGGAAGGCCAGGGAATTTCTTCAGAGCAGACCCCGCCCAACAGCCCGCCCGCC 1380
Qy      1369 GAGAGCTTCCGCTTCAGAGGAGACCCAC 1394
Db      1381 GCGAGCTTCAGGCTCGGGGTGCGGAC 1406
```

RESULT 2

```

US-09-552-950-5
; Sequence 5, Application US/09552950
; Patent No. 5541248
; GENERAL INFORMATION:
; APPLICANT: Oxford Biomedica (UK) Limited
; TITLE OF INVENTION: Anti-Viral Vectors
; FILE REFERENCE: 674524-2004
; CURRENT APPLICATION NUMBER: US/09/552,950
; CURRENT FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 9772
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: pSYNG?
US-09-552-950-5
```

```

Query Match      63.5%; Score 939.8; DB 4; Length 9772;
Best Local Similarity 78.6%; Pred. No. 1.3e-133;
Matches 1182; Conservative 0; Mismatches 297; Indels 24; Gaps 4;

Qy      1 ATGGGCGCGCGCGCCAGCATCTCTGCGCGCGCGCAAGCTGGACGCTGGAGCGCATCCGC 60
Db      1108 ATGGGCGCGCGCGCCAGCGTGTCTGCGCGCGCGAGCTGGACCGCTGGAGAGATCCGC 1167
Qy      61 CTGCGCGCGCGCGCGCAAGAGTGTCTACATGATGAAGCACTGGTGTGGGCCAGCGCGAG 120
Db      1168 CTGCGCGCGCGCGCGCAAAAGAGTCAAGCTGAAGCACTGTGTGGGCCAGCGCGAA 1227
Qy      121 CTGGAGAGTTCGCCCTGAACCCCGCGCTGTCTGGAGACCGAGCGGCTGCAAGCAGATC 180
Db      1228 CTGGAGCGCTTCGCCGTGAACCCCGCGCTCTCTGGAGACCGAGCGGTGCCCCAGATC 1287
Qy      181 ATCCGCGAGCTGCACCCCGCCCTGCAGACCGCGCAGCGAGAGCTGAAGAGCCTGTTCAAC 240
Db      1288 CTCGGCCAACTGCAGCCCGCCCTGCAACCCCGCGCAGCGAGAGCTGCGCAGCCTGTACAA 1347
Qy      241 ACCGTGGCCACCCCTGTACTGCGTGCACGAGAAAGATCGAGTCCGCGCACACCAAGGAGCC 300
Db      1348 ACCGTGGCCACCGTGTACTGCGTCCACCGCGCTCGAAATCAAGGATACGNAAGAGGCC 1407
Qy      301 CTGGACAAGATCGAGGAGGAGGAGGAGTGAAGTGCAGAGCAAGATCCAGCAGGCC - - - - - 354
Db      1408 CTGGATAAAATCGAAGAGGAGGAGGAGTGAAGAGCAAAAGAGGCGCCACAGCGCGCGCG 1467
Qy      355 GAGGCGCGCGCACAGGCGAAGGTGAGCGCGAGAACTACCCCATCGTGCAGAACTGCAGGCG 414
Db      1468 GACACCGGACACAGCAACCGAGGTCAAGCGCAAGAACTACCCCATCGTGCAGAACTCCAGGG 1527
Qy      415 CAGATGTTGCACACCGGCACTCAGCCCGCCCGACCCCTGAAGCGCTGGGTGAAGGTGATCGAG 474
Db      1528 CAGATGTTGCACACCGGCACTCTCCCGCCCGCACCGCTGAACCGCTGGGTGAAGGTGTTGGA 1587
Qy      475 GAGAAGCGCTTCAGCCCGCGAGTGTATCCCATGTTTACCGCGCTGAGCGAGCGCGCCACC 534
Db      1588 GAGAAGCGCTTTAGCCCGCGAGGTGATACCCATGTTTCTCAGCCCTGTTCAGAGGGAGCCACC 1647
Qy      535 CCCCAGGACCTGAACACGATGTTGAACACCGTGGCGGCCACCGAGCGGCGCCCAIGCAGATG 594
Db      1648 CCCCAGATCTGAACACCATGCTCAACACAGTGGGGGACACCCAGCGCCCGCCATGCAATG 1707
Qy      595 CTGAAGGACACCATCAACAGAGGAGGCGCGCGAGTGGGAGCGCGTGCACCCCGCGGCGCC 654
Db      1708 CTGAAGGAGACCATCAATGAGGAGGCTGCGGAATGGGATCGTGTGCATCGCGGTGCACGCA 1767
Qy      655 GGCCCCATGCCCGCCCGCGCAGATGCGGAGCGCCCGCGGCGGAGCGGAGCATCGCGCGCACCC 714
Db      1768 GGGCCCAATCGCACCGCGCGCAGATGCGGTGAGCCACCGGGGCTCAGACATCGCGCGGAGCAG 1827
Qy      715 AGCACCCCTGCAGGAGCAGATCGCCTGGATGACCAAGCAACCCCGCCCATCCCGCTGGGCGAC 774
```

```
Db 1828 AGTACCCTTCAGGAACAGATCGCGCTGGATGACCAACAACCCACCCCA1CCCGGTGGGAGAA 1887
QY 775 ATCTACAAAGCGGTGGATCATCTCTGGGCTGAGCAAGATGTTGGGATGTACAGCCCCCGT 834
Db 1888 ATCTACAAACGCTGGATCATCTCTGGGCTGAGCAAGATGTTGGGATGTATAGCCCTACG 1947
QY 835 AGCATCCTGGACATCAAGCAGGCGCCCAAGAGCCCTTCGCGGACATAGTGGACCCCTTC 894
Db 1948 AGCATCCTGGACATCCGCAAGGCGCCCAAGAGCCCTTCGCGGACATAGTGGACCCCTTC 2007
QY 895 TTCAAGACCCCTGGGCGCGAGGAGAGACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 954
Db 2008 TACAAAACGCTTCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2067
QY 955 CTGCTGGTGCAGAACGCCAACCCCGACTGCAAGAGCCATCTCTGGGCTCTCGGCGCCGCG 1014
Db 2068 CTGCTGGTGCAGAACGCCAACCCCGACTGCAAGAGCCATCTCTGGGCTCTCGGCGCCGCG 2127
QY 1015 GCCAGCCTGGAGGAGATGATGACCGCTGCGAGGCGCTGGGCGGCGCCAGCCACAGAGCC 1074
Db 2128 GCTACCCTAGAGGAATGATGACCGCTGTCAGGAGGAGTGGGCGGAGCCCGGACAGAGCA 2187
QY 1075 CGGCTGCTGGCGGCGGATGAGCGAGGCGCAACA-----CCAGCGTGATGATGAGAG 1124
Db 2188 CGGCTGCTGGCTGAGGCCATGAGCGAGGTGACCAAC1CGGCTACCATCA1GATCGAGCG 2247
QY 1129 AGCAAC1TCAAGGGCGCCCGCGGCA1CGTCAAGTGTCTTCAACTGGGCGCAAGAGGCGCAC 1188
Db 2248 GCGAAC1TTCGGAACCAACGCAAGATGTCAGAGTGTCTTCAACTGGGCGCAAGAGGCGCAC 2307
QY 1189 ATCGCGCGCAACTGCGCGCGCGCGCGCGCAAGAGGCGTGTCTGGAGAGTGGGCGAAGAGGG 1249
Db 2308 ACAGCGCGCAACTGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2367
QY 1249 CACGAGATGAAGGACTGCACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1308
Db 2368 CACCAATGAAGATGTACTGAGAGACAGCGCTAATTTTATAGGAAGATCTGGCCCTCC 2427
QY 1309 CACAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1368
Db 2428 CACAAGGGAAGCGCAGGGAATTTCTTCAGAGCAGACCAAGAGCGCAACAGCCCGCAGAA 2487
QY 1369 GAGAGCTTCGGCTCGAGGA-----GACCACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1416
Db 2488 GAGAGCTTCAGGTTTGGGGAAGAGACACAACACTCCCTCTCAGAGCAGGAGCGCGATAGAC 2547
QY 1417 AAGGACCGCGGAGACCCCTGACAGGCTGAAGAGCGCTGTTCGGCAAGCAGCCCGCGAGCGAG 1476
Db 2548 AAGGAAC1G1ATCCTT1AGCTTCCCTCAGATCACTCTTTTGGCAGGACCCCTCGTCACAA 2607
QY 1477 TAA 1479
Db 2608 TAA 2610

RESULT 3
US-09-184-418C-100
; Sequence 100, Application: us/09184418c
; Patent No. 6492110
; GENERAL INFORMATION:
; APPLICANT: Hahn, Beatrice
; APPLICANT: Gao, Feng
; APPLICANT: Shaw, George
; TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
; TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS TYPE 1
; FILE REFERENCE: D6287
; CURRENT APPLICATION NUMBER: US/09/184,418C
; CURRENT FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 112
; SEQ ID NO 100
; LENGTH: 1476
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
```

```
; FEATURE:
; OTHER INFORMATION: isolate=941N476.104; gene=gag
US-09-184-418C-100

Query Match 49.9%; Score 738.2; DB 4; Length 1476;
Best Local Similarity 69.2%; Pred. No. 2.7e-103;
Matches 1023; Conservative 0; Mismatches 453; Indels 3; Gaps 1;

QY 1 ATGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
Db 1 ATGGGTGCGAGAGCGCTCAATATTAAAGAGGGGGAAATATAGATAGATGAGGAAATTCGG 60
QY 61 CTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Db 61 TTAAGGCCAGGGGAAAGAAACATTATATGATAAAACACTTAGTATGCGCAAGCAGGGAG 120
QY 121 CTGGGAGAGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
Db 121 CTGGAAGAATTTCGCGCTTAACCCCTGGCGCTTTAGAGACGTCAGACGGATGTAAACAATA 180
QY 181 ATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
Db 181 ATAAACAGCTACATCCAGCTCTTAAGACAGGAAACAGAGGAACATTAGTCAATTATTCAAC 240
QY 241 ACCGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
Db 241 ACAGTAGCAACTCTCTATTGTGTACATGCGAGGGATAGAGGTACGAGACACCAAGGAGCC 300
QY 301 CTGGACACATCGAGGAGCGAGACAGAAAGTGTCCAGAGAGATCCAGAGCGCGCGAGGCC 360
Db 301 TTAGACAAGATAGAGGAGAAACAACAAGTCAAGTCAAGCAAAAATACAGAGGCGCAAGAG 360
QY 361 GCGCACAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
Db 361 GCTGAC---GGAAAGGTCACTCAAAATTATCTTATAGTCAAAATCTCCAAGGGCAATG 417
QY 421 GTGACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
Db 418 GTACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 477
QY 481 GCCTTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
Db 478 GCTTTTAGCGCGAGGTTAATACCCATGTTACAGCATATPCAGAGGAGCGCGCGCGCGCT 537
QY 541 GACCTGAACACGATGTTGAACACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
Db 538 GATTTAACACCGATGTTAATACAGTGGGGGAGCATCAAGAGCGCATGCAAAATGTTAAA 597
QY 601 GACACCATCAACGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
Db 598 GATACCATCAACGAGGAGCGTGGGAATGGGATAGATTACATCCAGTACATCCAGGCGCT 657
QY 661 ATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
Db 658 AATCCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 717
QY 721 CTGCGAGGAGGATCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780
Db 718 CTTCAGGAACAATAAGCATGCGATGACAGGTTAACCCCACTATCCAGTGGGAGGACATCT 777
QY 761 AAGCGGTGATCATCTCTGGCGCTGAACAAGATGCTGGGATGTACAGCCCGCGCGCGCG 840
Db 778 AAAAGATGGATAATTCTGGGTTTAAATAAAATAGTAAGATGTATAGCCCTGTGAGCAAT 837
QY 841 CTGGACATCAAGCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
Db 838 TTGGACATAAGACAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 897
QY 901 ACCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 960
Db 898 ACTTTAAGAGCTGAACAAGCTACACAAGAGAAAGTAAAGGTTGATGACAGACACCTTGT 957
QY 961 GTGCAAGACGCGCAACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020
```

Db 958 GTCCAAATGCAACCCAGATTGTAAGACCAATTAAAGAGCAATAGGACCCAGGGCTTCA 1117
Qy 1021 CTGGAGGAGATGATGACCCCTGCGAGGGCGTGGGGGGCCCAAGCCACAAAGGGCGGGS 1080
Db 1018 TTAGAAGAAATGGTACAGCATGTCAAGGAGTGGGAGGACCTAGCCACAAAGGAGTS 1077
Qy 1081 CTGGCCGAGGGGATGAGCCAGGCCAACACCCAGGCGTGAATGATGCAAGAGCAACTTCAA 1140
Db 1078 TTGGCTGAGGCAATGAGCCATTCACATGTACATAAATGATGCAAGAGGCAATTTTAA 1137
Qy 1141 GGCCCGCCGCGCATCGTCAAGTGTCTCAACTGCGCAAGGAGGGCCACATCGCCGCAAC 1200
Db 1138 GGCCCTAAAGAAATGTTAAATGCTCAACTGTGGCAAGGAGGACACCAATGAAA 1197
Qy 1201 TGCCCGCCCGCCCGCAAGGAGGGTGTGGAAGTCCGGCAAGGAGGGCCACCAAGTGAAG 1260
Db 1198 TGACGGGCCCCCTAGAAAAAGAGGCTGTGGAAATGTGGCAAGGAGGACACCAATGAAA 1257
Qy 1261 GACTGCACCGAGCGCCAGGCCAACTTCTGGGCAAGATCTGGCCAGCCACAAAGGCCCGC 1320
Db 1258 GACTGTACTGAGAGCGAGGCTAATTTTATAGGAAATTTGGCCTTCCACAAAGGGGAGG 1317
Qy 1321 CCGGCAAACTTCTGCGAGAGCCCGCCGAGCCCAAGCCCGCCCGCGGAGGCTTCGC 1380
Db 1318 CCAGGAAATTCCTTCAAAACAGGCGCAGAGCCCAAGCCCGCCCGGAGGCTTCAGG 1377
Qy 1381 TTGAGGAGAGACCCACCCCGCCGAGAGCAGGAGGAGGAGCCGAGACCCCTGACCAAGC 1440
Db 1378 TTCAAGGAGACAAACCCCGCTCCGAGCAGGAGTCCGAAAGACAGGGAACCCCTTAACCTCC 1437
Qy 1441 CTGAAGAGCCCTGTTGGCAACGACCCCTGAGCCAGTAA 1479
Db 1438 CTCAAATCACTCTTGGCAGCGACCCCTGTCTCAATAA 1476

RESULT 4

US-09-184-418C-11
; Sequence 11, Application US/09184418C
; Patent No. 6492110
; GENERAL INFORMATION:
; APPLICANT: Hahn, Beatrice
; APPLICANT: Gao, Feng
; APPLICANT: Shaw, George
; TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
; TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS TYPE 1
; FILE REFERENCE: D6287
; CURRENT APPLICATION NUMBER: US/09/184,418C
; CURRENT FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 112
; SEQ ID NO 11
; LENGTH: 8959
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; OTHER INFORMATION: isolate="94IN476.104; 138.1613;" "aaq";
; OTHER INFORMATION: 1418.4428;"poi"; 4361.4939;"vif"; 4879.5169;"vpr";
; OTHER INFORMATION: 5150.7782;"tat"; 5289.7939;"rev"; 5378.5638;"vpu";
; OTHER INFORMATION: 5556.8129;"env"; 8131.8754;"nef"
US-09-184-418C-11

Query Match 49.98; Score 738.2; DB 4; Length 8959;
Best Local Similarity 69.28; Pred. No. 2.6e-103;
Matches 1023; Conservative 0; Mismatches 453; Indels 3; Gaps 1;
Qy 1 ATGGGCGCGCGCCAGCATCCCTGGCGGGCGGCAAGCTGGACGCCCTGGAGCGCATCGGC 60
Db 138 ATGGGTGGAGAGCGTCAATATTAAAGGGGGGAAATTAGATAGATGGGAAATAATTCGG 197
Qy 61 CTGCGCCCGCGCGGCAAGAGTGTACATGATGAAGCACCTGGTGGGCCAGCCCGGAG 120
Db 198 TTAAGGCCAGGGGGAAGAAACATTTATATGATAAAACACTTAGTATGGGCAAGCGGGAG 257

Qy 121 CTGGAGAACTTCGCCCTGAACCCCGGCCCTGCTGGAGACACAGCGAGGGCTGCAAGCAGATC 180
Db 258 CTGGAAGATTTGGCTTAACCTTGGCCCTTTAGAGACCTCAGACGGATTAACAATA 317
Qy 181 ATCCGCCAGCTGCACCCCGCCCTGACAGACCGGCGAGGAGCCTGAAGAGCCTGTTCAAC 240
Db 318 ATAAACAGCTACATCCAGCTCTTAAGACAGGACACAGAGCAACTTAGTTCATTATCAAC 377
Qy 241 ACCGIGGCCACCCCTGTACTGCTGCACGACAGAGATCGAGGTCGCGGACACCAAGGAGGCC 300
Db 378 ACAGTAGCAACTCTCTATTGTGTACATGCGAGGGATAGAGGTACGAGACACCAAGGAGGCC 437
Qy 301 CTGGACAAGATCGAGGAGGAGCAGAAACAGTSCCAGCAGAAAGATCCAGAGGCCGAGGCC 360
Db 438 ITAGACAAGATAGAGGAAGAACAAACAAAAGTCAGCAAAAATACAGCAGGCAAAAGAG 497
Qy 361 GCCGACAAGGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAAACCTGCAGGGCCAGATG 420
Db 498 GCTGAC--GGAAGGTCAGTCAAAATTTCTCTATAGTGCAAAATCTCCAAGGGCAATG 554
Qy 421 GTGCACCGGCCCATCAGCCCGCCGACCCCTGAACGCCCTGGTGAAGGIGATCGAGGAGAG 480
Db 555 GTACACCGGCCCTATCACCTAGAACCTTGAATGCGTGGTAAAGTAAATAGAGGAGAG 614
Qy 481 GCCITCAGCCCGAGGTGATCCCATGTTCAACCCCTGAGCGAGGGCGCCACCCCGCAG 540
Db 615 GCTTTAGCCCGAGGTATACCCATGTTACAGCATTATCAGAGGAGGCCACCCCTCT 674
Qy 541 GACCTGAACACCGATGTTGAACACCGTGGGCGGCCACCCAGGCCCATCGAGATGCTGAAG 600
Db 675 GATTTAAACACCATGTTAATAACAGTGGGGGGACATCAAGCAGCCATGCCAATGTTAAA 734
Qy 601 GACACCATCAACGAGGAGGGCGCCGAGTGGGACCCCGTGCACCCCGTGCACCGCGGCC 660
Db 735 GATACCATCAACGAGGAGGCTCGGAATGGGATAGATTACATCCAGTACATGCGAGGGCCT 794
Qy 661 ATCGCCCGCCCGCAGATGCGGAGCCCGCGGCGAGCAGATCGCGGCGCACATCGCGGCGCACCCAGCAC 720
Db 795 AATCCACCGAGCCAGATGAGAGAACCAAGGGGAAGTATATAGCAGGAACCTACTAGTACC 854
Qy 721 CTGCAGGAGCAGATCGCTGGATGACCAAGCAACCCCGCCATCCCGCTGGGCGACATCIAC 780
Db 855 CTCAGGAACRAATAGCATGGATGACAGGTAAACCCACCTATTCAGTGGGAGACAICTAT 914
Qy 781 AAGCGGTGGATCATCTCTGGGCTGAACAAGATCGTGGGATGTACAGCCCGCTGAGCATC 840
Db 915 AAAAGATGGATAATTCTGGGTTAAATAAATAGTAAGTAATCTATAGCCCTGTCAGCAT 974
Qy 841 CTGGACATCAACGAGGCGCCCAAGGAGCCCTTCGCGGACATCGTGGACCGCTTCTTCAAG 900
Db 975 TTGGACATAGACAAAGGGCCCAAGGAACCCCTTAGAGACTATGTAGACCGGTCTTAA 1034
Qy 901 ACCCTGGCGCGGAGCAGACACCCAGGAGGTGAAGAATCGATGACCGACACCCCTGCTG 960
Db 1035 ACITTAAGAGCTGAACAAGCTACACAGAGTAAGGTAAAGGTGATGACAGACACCTGTTG 1094
Qy 961 GTGCAGAACGCCAACCCCGACATGCAAGACCATCTCGGCGCTCTCGGCCCGCGCCAGC 1020
Db 1095 GTCCAAAATGCCAACCCAGATGTAAAGACCATTTAAGAGCATTAGGACCGGGCTTCA 1154
Qy 1021 CTGGAGGAGATGATGACCCCTGCCAGGGCGGTGGCGGCCAGCCACAAAGGCCCGGCTG 1080
Db 1155 TTAGAAGAAATGGTGACAGCATGTCAAGGAGTGGGAGGACTAGCCACAAAGCAAGAGTG 1214
Qy 1081 CTGGCCGAGCGGATGAGCCAGGCCAACACAGCGCTGATGATGACAGAGAGCAACTTCAAG 1140
Db 1215 TTGGCTGAGGCAATGAGCCCAATCACAATAGTAACATAATGATGACAGAGAGCAATTTTAA 1274
Qy 1141 GGCCCGCCCGCGCATCGTCAAGTGTCTCAACTGGGCGCAAGGAGGCCACATCGCCCGCAAC 1200
Db 1275 GGCCCTAAAGAAATGTTAAATGCTTCAACTGTGGCAAGGAGGCGCACATAGCCAGAAAT 1334
Qy 1201 TGCCGCGCCCGCCGCAAGAGGGCTGCTGGAAGTGGGCAAGGAGGGCCACAGATGAAG 1260

```
Db 1335 TGCAGGGCCCTAGAAAAAGAGGCTGTTGGAAATGTGGCAAGAGCACCAATGA 1394
QY 1261 GACTGCAGCGAGCGCCAGGCCAACTTCCTGGCAAGATCTGGCCAGCCACAAGGGCCGC 1320
Db 1395 GACTGTACIGAGAGCGCAGGCTAAATTTTTFAGGAAATTTTGCCCTTCCCAAGGGGAGG 1454
QY 1321 CCGGGCAACITTCCTGCAGAGCGCGCCGAGGCCACCGCCCGCCCGCGGAGAGCTTCGGC 1380
Db 1455 CCAGGGAATTCCTTCAAAACAGCGCCAGAGCCCAACAGCCCGCCAGCAGAGAGCTTCAGG 1514
QY 1381 TTCAGGAGACCAACCCCGCGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440
Db 1515 TTCAGGAGACCAACCCCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1574
QY 1441 CTGAAGAGCCTGTTCGGCAACGACCCCTTGAGGCCAGTAA 1479
Db 1575 CTCAAATCACTCTTTGGCAGCGACCCCTTGTCCTCAATAA 1613

RESULT 5
US-09-184-418C-73
; Sequence 73, Application US/09184418C
; Patent No. 6492110
; GENERAL INFORMATION:
; APPLICANT: Hahn, Beatrice
; APPLICANT: Gao, Feng
; APPLICANT: Shaw, George
; TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
; TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS TYPE 1
; FILE REFERENCE: D6287
; CURRENT APPLICATION NUMBER: US/09/184,418C
; CURRENT FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 112
; SEQ ID NO 73
; LENGTH: 1485
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; OTHER INFORMATION: isolate=962M651.8; gene=gag
US-09-184-418C-73

Query Match 49.3%; Score 729.8; Dh 4; Length 1485;
Best Local Similarity 69.2%; Pred. No. 4,9e-102;
Matches 1027; Conservative 0; Mismatches 452; Indels 6; Gaps 2;

QY 1 ATGGGGCGCCCGCCAGCAICTCTGCGCGGCGGCAAGCTGGAGCCCTGGGAGCGCAICCGC 60
Db 1 ATGGGIGGAGAGCGTCAATATTAAGAGGGGGAATTAGATAATTTGGGAAAAATTAGG 60
QY 61 CTGCGCCCGCGCGGCAAGAGTGCTACATGATGAAGCACCTGGTGGTGGCGAGCGCGGAG 120
Db 61 CTAAGGCCAGGGGAAAGAACGCTATATGATAAACACCTAGTATGGGCAAGCAGGGAG 120
QY 121 CTGGAGAAGTTCCGCTGAACCCCGGCTGCTGGAGACCGAGGAGGCTGCAAGCAGATC 180
Db 121 CTGGAAGATTTCGCTTAACCTTGGCTTTTGAACAATCAGAAGCGCTGTAACAATA 180
QY 181 ATCCGCGCAGCTGCACCCCGCCCTGCAGACCGCGCAGGAGGAGCTGAAGAGCGCTTCAAC 240
Db 181 ATGAACAGCTACACCCGCTTTCAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
QY 241 ACCGTGGCCACCTGTACTGCTGCGACGAGAGATCGAGGTCCGCGACACCAAGGAGGAG 300
Db 241 ACAGTAGCAACTCTCTATCTGTGTACATGAAGGGGTAGAGGTACGAGACACCAAGGAGCC 300
QY 301 CTGGACAAGATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 357
Db 301 TTACACAGGATAGAGGAAGAGCAACAAATTCAGCAAAAATATCAGCAAAAATACAGCAAAAACAG 360
QY 358 GCGCGCGACAGGGCAAGGTGAGCCAGAACTACCCCATCTGTCAGAACTGTCAGAGGGCGAG 417
Db 361 CAAGCGGCTGACGGAAAGGTGAGTCAAAAATATCTCTATAGTGCAGATCTCCCAAGGGCA 420
```

```
QY 418 ATGGTGACCAGGCCATCAGCCCCCGCCACCTTGAAAGCCCTGGGTGAAGGTGAICGAGGAG 477
Db 421 ATGGTACACCAGAACTATCACTAGAACTTTGAATGCAITGGTAAAGTAAIAGAAGAA 480
QY 478 AAGGCGCTTCAGCCCGGAGGTGATCCCATGTTTCAAGCCCTTGAGCGAGGGCGCCACCC 537
Db 481 AAGCGTTTTCAGCCCGGAGGTGATCCCATGTTTACAGCAITATCAGAAGGAGCCACCCCA 540
QY 538 CAGGACCTGAAACACGATGTTGAACACCGTGGCGGCGCCACCAAGCCCGCATGCAGATGCTG 597
Db 541 CAAGATTTAAACACCATGTTAAATACAGTGGGGGACATCAAGCAGCCATGCAATGTA 500
QY 598 AAGGACACCATCAACGAGGAGGCGCGCGAGTGGGACCGCTGCACCCCGTGCACGCGGCG 657
Db 601 AAGGATACCTATCAATGAGGAGGCTGCAGATGGGATAGATTACATCCAGTGCATGAGGG 660
QY 656 CCGATCGCCCGCGCGAGATGCGGAGCGCCCGCGGAGCGGAGCATCGCGGCGCACCAAGC 717
Db 661 CCTATTGCACCGGCGCAATGAGAGAACCAAGGGGAGTGATATAGCAGGAACCTACTAGT 720
QY 718 ACCCTGAGGAGCAGATCGCTGGATGACCAAGCAACCCCGCATCCCGTGGCGCACATC 777
Db 721 ACCCTCAAGCAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 778 TACAAGCGGTGATCATCTCTGGCTGAAACAAGATGCTGGGATGTACAGCCCGGTGAGC 837
Db 781 TATAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 838 ATCTTGACATCAACGAGGCGCCCAAGGAGCGCTTCGCGGACTACGTGGACCGCTTCCTC 897
Db 841 ATTTTGGACATAAAACAAGGGCCCAAGGAGCCCTTTAGAGACTATCTAGACCGGTTCTC 900
QY 898 AAGACCTGCGCGCGCGAGCAGACCCAGGAGGTGAAGAACTGGATGACCGGACACCTG 957
Db 901 AAAACITTAAGAGCTGAACAGGCTACACAAGAGTAAATAATTTGGATGACAGACACCTG 960
QY 958 CIGGTGCAGAACGCCAACCCCGACTGCAAGACCATCTCGCGCTCTCGCGCGCGCGCGC 1017
Db 961 TTGTTCCAAAATGCAAAACCCAGATTGCAAGGCCATTTTAAAGCATIAGGACCGAGGGCT 1020
QY 1018 AGCCTGAGGAGATGATGACCGCTGCGAGGGCGTGGCGGCGCCCGCCAGCCACCAAGCGCGC 1077
Db 1021 ACATTAGAAGAAATGATGACAGCATGTCAAGGAGTGGGAGGAGCTAGCCACAAGCAAGA 1080
QY 1078 GTGTCGCGCGGCGGATGAGCCAGGCGCAACA--CCAGCGTGTATGATGCAGAAGAGCAAC 1134
Db 1081 GGTTCGCTGAGGCAATGAGCCAAACAATAATAGTGTAAACATACTGTATGCAAGAAAGCAAT 1140
QY 1135 TTCAAGGCGCCCGCGCGCATCTGTCAGTGTCTCAAGTGTCTCAAGTGGCAAGGAGGCGCACATCGCC 1194
Db 1141 TTTAAAGGAAATAAAAGAAATGTTAAATGTTTAACTGTGTAAAGAAAGGCGCACATAGCC 1200
QY 1195 CGCAACTGCGCGCGCGCGCGCAAGAGGGCTGTGGAAGTGGCGCAAGGAGGCGCACACAG 1254
Db 1201 AGAAATGCGAGGCGCCCTAGGAAAAGGGCTGTTGGAATGTGGAAGGAGGAGGAGCAACAA 1260
QY 1255 ATGAAGGACTGCACCGAGCGCCAGGCAAACTTCCTGGCAAGATCTGGCGCCAGCCACAAG 1314
Db 1261 ATGAAGGACTGTACTGAGAGGCGAGGCTAATTTTATAGGAAATTTGGCCTTCCCAAG 1320
QY 1315 GCGCGCGCGCGCGCAACTTCCTGCAGAGCGCGCGCGAGCCCGCCCGCGCGCGCGCGAGC 1374
Db 1321 GGAAGCGCAGGGAATTCCTTCAGAACAGCGCCAGAGCCCAACAGCCCGCCAGCAGAGC 1380
QY 1375 TTCGCTTCGAGGAGACCAACCCCGCGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGC 1434
Db 1381 TTCAGGTTGAGGAGACCAACCCCGCGCTCCGAAGCAGGAGGTGGAAGAGCAGGGAGCCTTA 1440
QY 1435 ACCAGGCTGAAGAGCGCTGTTGCGCAAGGAGCCCGCTGAGCCAGTAA 1479
Db 1441 ACTTCCTCAATCACTCTTTTGGCAGGAGCCCGCTTGTCTCAATAA 1485
```



```
/ LENGTH: 1486
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; OTHER INFORMATION: isolate-92RW009; gene-gag
US-09-184-418C-38

Query Match      48.1%; Score 710.8; DB 4; Length 1486;
Best Local Similarity 68.8%; Pred. No. 3.5e-99;
Matches 1022; Conservative 0; Mismatches 457; Indels 7; Gaps 3;

QY 1 ATGGGCGCCCGCCAGCCATCCCTGCGCGGGCAAGCTGACAGCCCTGGGAGCGCA-TCGG 59
Db 1 ATGGGTGGAGAGCGTCAATATTAAGAGGCGGAAATTAAGTGCCTGGGAAAAATTAAG 60

QY 60 CCTGCGCCCGCGCGCAAGAGTGTACATGATGAAGCACTGTGTGGGCCAGCGCA 119
Db 61 TTAAAGCCAGGGGGAAGAAACATATATGATGAAACACCTAGTAGTGGCAAGCAGGGA 120

QY 120 GCTGGAGAAGTTCGCCCTGAACCCCGGCTGCTGGAGACAGCGAGGGTSCAAGCAGAT 179
Db 121 GCTGGAAGATTTCACCTTAACCCCTGACCTTTAGACACACCAAGAGGCTGTAAACAAAT 180

QY 180 CATCCGCCAGTCACCCCGGCTGACAGCGCGGCGAGAGTGCAGAGAGCTGAAGAGCTGTAA 239
Db 181 AATGACACAGCTGCANCCAGCTCTTCAGACAGGAAACAGATGAACCTTAGTGCATTATATAA 240

QY 240 CACCGTGGCCACCCCTGTACTGTGTCACGAGAAAGATCGAGTCCGCGACACCAAGGAGCG 299
Db 241 TACAGTAGCAACCCCTCTATTGTGTACATCAAAAGATAGATGTAAAGACACCAAGGAGGCG 300

QY 300 CCTGGACAGATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 359
Db 301 CTTAGACAAGATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360

QY 360 CGCCGACAAGGCAAGGTGAGCGCAGAACTACCCCATCGTGCAGAACTTCGAGCGGCGCAGAT 419
Db 361 AGCTGACAAAGGAAAAAGTCACTCAAAAATTACCCCTATAGTGCAAAATGCAAGAGGGCAAT 420

QY 420 GGTGCACACGAGGCATCAGCCCGCGCACCCCTCAACCGCTGCTGCTCAAGGTGATCGAGGAGAA 479
Db 421 GGTACACACGAGGCATATCACCCTAGAACTTTGAATGCGTGGTAAAGTATAGAGGAGAA 480

QY 480 GGCCTTCAGCCCGGAGGTGATCCCAAGTTCACCGCCCTGAGCGAGGGCGCCACCCCGCA 539
Db 481 GGCCTTTAGCCAAAGAGGTATACCCCATGTTACAGCATTTATCAGAAAGGAGCGACCCGACA 540

QY 540 GGACCTGAACAGATGTTGAACACCGTGGCGGCGCACCGAGCGCGCCATGACAGATGCTGAA 599
Db 541 AGATTAAACACCATGCTAAATACAGTGGGGGACATCAACAGCCCATGCAAAATGCTAAA 600

QY 600 GGACACCATCAACAGGAGGAGGCGCGCGAGTGGACCGCGTGCACCCCGCTGACCGCGCGCG 659
Db 601 AGATACATCAATGAGGAGGAGGCTGCAGAGTGGGATAGGTACATCCAGTGCAGGCGAGGGCG 660

QY 660 CATCGCCCGCGCCAGATGCGCGAGCCCGCGGCGAGCGACATGCGCGGACACCCAGGAC 719
Db 661 TGTTCGCCAGCGCCAGATTAAGAGAACCAAGGGGAAGTGACATAGCAGGAAGTACTAGTAG 720

QY 720 CCTGCAGGAGCAGATCGCCTGATGACGAGCAACCCCGCCATCCCGTGGCGGACATCTA 779
Db 721 CTTTCAGGAACAAATAGCATGATGATGACAAATAACCCACCTATTCAGTGGGAGAAATTA 780

QY 780 CAAGCGGTGATCATCCTGGGCTGAACAAAGATCGTGGGATGTACAGCCCGCTGAGCAT 839
Db 781 TAAAGATGGATAAATCTGGGTTAAATAAAATAGTAAGATGTATAGCCCTGTGAGCAT 840

QY 840 CTTGGACATCAACAGGCGCCCAAGGAGCCCTTCGGGACTACGTGGACCCCTTCITCAA 899
Db 841 ATTGGACATAAAACAAGGCGCCAAAGGAACCTTTAGAGACTATGTAGACCGGTTCTTTAA 900

QY 900 GACCCTGCGCGCGGAGCAGACGACCCAGGAGGTGAAGAAGTGGATGACCGGACACCCCT 959
Db 901 AACCTTAAGAGCGCGCAAGAGCTTCACAAGATGTAAAAAATTGGATGACAGATACCTTGT 960

QY 960 GGTGCAGAAACGCCAACCCCGACTCCCAAGACCATCTCTGCGGCTCTCGGCCCCGGCGCCAG 1019
Db 961 AGTCCAAAATGCGAAACCCAGATTGTAAAGACCATTTTAAGACCATTAAGGCCAGGGCTTC 1020

QY 1020 CCTGGAGGAGATGATGACCGCCCTGCCAGGGCGGTGGGCGGCCCGCCAGCCACAAGGCCCGGT 1079
Db 1021 ATTAGAAGAAATGATGACAGCATGCCAGGGAGTGGGAGGACCCCGGCCATAAAGCAAGGGT 1080

QY 1080 GCTGCGCGGAGGCGATGAGCCAGG---CCAACACCAAGCGTGTATGATGACAGAGCAACTT 1136
Db 1081 TTTGGCTGAAGCAATGAGCCCAAGTACAAACAACCAACATATATGATGACAGAGGCAATTT 1140

QY 1137 CAAGGCCCCCGCGGCATCGTCAAGTCTTCAACTGCGGCAAGAGGGGCCACATCGGCCCG 1196
Db 1141 TAAGGCCAGAGAAGAAATTAATTAAGTGTTCCTCAACTGTGGCAAAAGAGGACACCTAGCCAG 1200

QY 1197 CAACTGCGCGCGCCCGCGCAAGAGGGCTGCTGGAAGTGGGCAAGGAGGGCCACCCAGAT 1256
Db 1201 AAATTGCGAGGGCCCTAGAAAAAAGGCGCTGTTGGAATGTGGAAGGAGGGACACCAAT 1260

QY 1257 GAAGGACTGCACCGAGCGCGCGCAACTTCCTGGCAAGATCTGGCCCGAGCCACCAAGGG 1316
Db 1261 GAAGGACTGCCTGAGAGACAGGCTAATTTTTTASGSAATTTTGGCTTCCCAACAAGGG 1320

QY 1317 CCGCCCGCGCAACTTCCTGACAGAGCGCGCGGAGCCCGGAGCCCGCCCGCGGAGGCTT 1376
Db 1321 GAGGCGAGGAAATTTTCCCGAGAGGAGACTGGAGCCACAGCCCGCCAGGAGGAACTT 1380

QY 1377 CCGCTT---CGAGGAGACCCCGCGCGCGCGGAGCCCGGAGGAGGAGGAGGAGGAGCCCT 1433
Db 1381 TGAATGGGGGAGAGATAGCCTCTGCTGTGAACAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440

QY 1434 GACCAAGCTGAAGAGCTGTTCGGCAACGACCCCTGAGCCAGTAA 1479
Db 1441 AATTTCCCTCAATCCTCTTTGGCAAGGACCCCTTGTACAGTAA 1485
```

RESULT 6

```
US-09-184-418C-4
; Sequence 4. Application US/09184418C
; Patent No. 6492110
; GENERAL INFORMATION:
; APPLICANT: Hann, Beatrice
; APPLICANT: Gao, Feng
; APPLICANT: Shaw, George
; TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
; TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS TYPE 1
; FILE REFERENCE: D6287
; CURRENT APPLICATION NUMBER: US/09/184,418C
; CURRENT FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 112
; SEQ ID NO 4
; LENGTH: 8992
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; OTHER INFORMATION: isolate=92RW009; 139.1624:gag; 1690.4428:pol(N-terminus unce
; OTHER INFORMATION: 4373.4951:vif; 4891.5181:vpr; 5162.7801:tat; 5301.7958:rev;
; OTHER INFORMATION: 5403.5648:vpu; 5566.8148:env; 8150.8773:nef
US-09-184-418C-4
```

Query Match

```
Best Local Similarity 48.1%; Score 710.8; DB 4; Length 8992;
Matches 1022; Conservative 0; Mismatches 457; Indels 7; Gaps 3;
```

```
QY 1 ATGGGCGCCCGCGCCAGCATCTCTGCGCGGCGGCAAGCTGGACGCGCTGGAGCGCA-TCGG 59
Db 139 ATGGTGGGAGAGCGCTCAATATTTAAGAGGCGGGAATATTAGATGCTCTGGGAAAAATTAAG 198

QY 60 CTTGCGCCCGCGGCGCAAGAGTGTGTACATGATGAGGACCTGTGTGGGCGGAGCCGCGCA 119
Db 60 CTTGCGCCCGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 119
```

Db 199 TTAAGCCAGGGGAAAGAAACATATAATGATGAACACCTAGTAAGGCAAGCAGCGA 258
QY 120 GCTGGAGAACTTCGCCCTGAACCCCGGCTGCTGGAGACCAAGGCTGCAAGCAGAT 179
Db 259 CTTGGAAGATTGCACTTAACCTGACCTTTAGAGACACCAAGGCTGTAAACAAAT 318
QY 180 CATCCGCCAGTGCACCCCGCTGCAGACCGCGCAGCGAGAGCTGAGAGCCTGTGAA 259
Db 319 AATGAGACAGCTGCAACCAAGCTCTTCAGACAGAACAGATGAACCTTAGCTATTATAA 378
QY 240 CACCGTGGCCACCTGTACTGGTGCAGAGAGATCGAGTCCCGCAGACCAAGAGGC 299
Db 379 TACAGTAGCAACCTCTATTGTGTACATCAAAAGATAGATGTAAAGACACCAAGAGGC 438
QY 300 CTTGACAAAGATCGAGGAGGAGCAACAAGTCCAGCAGAGATCCAGAGGCCAGGC 359
Db 439 CITAGACAGATAGAGGAAGAACAAACAAAGTCAGCAAAACACAGCAGGCAGAGC 498
QY 360 CGCCGACAAAGGAGGTGAGCAGAACTACCCCATCTGTCGAACCTGCAGGGCCAGAT 419
Db 499 AGCTGACAAAGGAAAGTCAGTCAAAATTACCTATAGTGCAAAATGCACAGGGCAAT 558
QY 420 GTGCAACAGGCCATCAGCCCCCGCACCCCTGACGCTGGTGAAGGTGATCGAGGAA 479
Db 559 GTACACACCGCCATATCACCTAGAACCTTGTATGCTGGTGAAGTAATAGAGGAA 618
QY 480 GGCCTTCAGCCCGGAGGTGATCCCCATGTTACCCGCTGACGAGGGCGCACCCCCCA 539
Db 619 GGCCTTTAGCCAGAGGIAATACCCCTGTTTACAGCATTTATCAGAAGGAGCCACCCACA 678
QY 540 GGAAGCTGAACAGATGTTGAACACCGTGGGCGGCCACAGGCCGCTGACAGTGCIGAA 599
Db 679 AGATTAAACACCATGCTAAATACAGTGGGGGACATCAAGCAGCCATGCAATGCTAAA 738
QY 600 GGACACCATCAACAGGAGGCGCGCGAGTGGGACCGGCTGCACCCCTGACAGCCGCGC 659
Db 739 AGATACAAATAGAGGAGGCTGCAGAGTGGGATAGGTACATCCAGTGCAGGAGGCG 798
QY 660 CATGCCCCCGGACAGTGGCGAGCGCGCGGACGAGACATGGCGGACACCAAGCAG 719
Db 799 TGTGCGCCAGCCAGATNAGAGAACCAAGGGAAGTGACATAGCAGGAACCTACTAGTAC 858
QY 720 CTTGAGGAGCAGATGCTGGATGACACAGCAACCCCGCTGCGGCGGACATCTA 779
Db 859 CTTGAGGACAAATAGCATGGATGACAAATACCCCGCTATCCAGTGGGAGAAATTA 918
QY 780 CAAGCGGAGTATCCTGGGCTGGAACAAGATGCTGGGATGTACAGCCGCTGAGCAT 859
Db 919 TAAAGATGGATAATCTGGGTTAAATAAAATAGTAAGAAATGTATAGCCCTGTACAGAT 978
QY 840 CTTGACATCAAGCAGGCGCCCAAGCAGCCCTTCGCGGACTACCTGGACCGGTTCTCA 899
Db 979 ATTGACATAAACAAGGCGCAAGGAACCTTTAGAGACTATGTAGACCGGTTCTTAA 1038
QY 900 GACCTGCGGCGGAGCAGAGCAGCCAGGAGGTGAAGAATGGATACCCGACACCCCTGCT 959
Db 1039 AACCTTAAGAGCGCAACAGCTTCACAAGATGTAAATAATTTGATGACAGATACCTTGT 1098
QY 960 GTGCAAGACGCCAACCCCGACTGCAAGACCATCTGCGGCTCTCGGCCCGCGGCGAG 1019
Db 1099 AGTCAAAATGCGAAGCCAGATTTAAGACCATTTAAGAGCATTAGGGCCAGGCGCTTC 1158
QY 1020 CTTGAGGAGATGATGACCGCTGCGAGGCGGCTGCGGCGCCAGCCACAGCGCGCT 1079
Db 1159 ATTAGAGAAATGATGACAGCATGCCAGGAGTGGGAGGACCCGCGCATTAAGCAAGGT 1218
QY 1080 GCTGCGGAGGCGATGAGCCAGG---CCAACACCGCTGATGATGCAGAGAGCAACTT 1136
Db 1219 TTTGCTGAGCAATGAGCCAAAGTACANCAACCAACATATGATGCAGAGAGCAATTT 1278
QY 1137 CAAGGCCCCCGGCGCATGTCAGTGTCTCACTGCGGCAAGGAGGCGGCACATCGCCCG 1196
Db 1279 TAAGGCGCAGAGAGAAATATTAAAGTGTTCAACTGTGGCAAGAGGACACCTAGCCAG 1338

QY 1197 CAACTGCCGCGCGCCCGCAAGAAAGGCTGTGAAAGTGGGCAAGGAGGGCCACCAAGAT 1256
Db 1339 AAATTGAGGGCCCTAGAAAAAAGGCGTGTGAAATGIGGAAGGAGGGACACCAAT 1398
QY 1257 GAAGGACTGCACCGCAGCGCCAGGCAACTTCCTGGGCAAGATCTGCGCCAGCCACAAGG 1316
Db 1399 GAAAGACTGCATGAGAGACAGGCTAATTTTTAGGGAATAATTGGCTTCCAAACAAGG 1458
QY 1317 CCGCCCGCGCACTTCCTGAGAGCGCCCGCGAGCCACCGCCCGCGGAGAGCTT 1376
Db 1459 GAGGCGAGGAAATTTCCCCAGAGCAGACTGGAGCCAAACAGCCCGCAGCAGAGAACTT 1518
QY 1377 CCGCTT--CGAGGAGACCAACCCCGCGCAGAGGAGGAGGAGGAGGAGGAGGAGCCCT 1433
Db 1519 TGAATGGGGAGAGATAGCCTCTCTCTCTGAAACAGGAGCAGAAAGACAGGAACTTT 1578
QY 1434 GACCAGCCTGAAGAGCTGTTGCGCAACGACCCCTGAGCCAGTAA 1479
Db 1579 AATTTCCCTCAATCACTCTTTGGCAACGACCCCTTGTACAGTAA 1624

RESULT 9

US-09-184-418C-82

: Sequence 82, Application US/09184418C

: Patent No. 6492110

: GENERAL INFORMATION:

: APPLICANT: Hahn, Beatrice

: APPLICANT: Gao, Feng

: APPLICANT: Shaw, George

: TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN

: TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS TYPE 1

: FILE REFERENCE: D6287

: CURRENT APPLICATION NUMBER: US/09/184,418C

: CURRENT FILING DATE: 1999-11-02

: NUMBER OF SEQ ID NOS: 112

: SEQ ID NO 82

: LENGTH: 1496

: TYPE: DNA

: ORGANISM: Human immunodeficiency virus type 1

: FEATURE:

: OTHER INFORMATION: isolate=962M751.3; gene=gag

US-09-184-418C-82

Query Match 47.4%; Score 701.2; DB 4; Length 1496;

Best Local Similarity 68.8%; Pred. No. 9,7e-98;

Matches 1032; Conservative 0; Mismatches 443; Indels 25; Gaps 4;

QY 1 ATGGGCGCGCGCGCCAGCATCCTGCGGGCGGCAAGCTGGACGCTGGGAGCGCATCCGC 60
Db 1 ATGGGTGCGAGAGCGTCAATATTAAAGAGCGGAAATTAGATGAATGGGAAAGAAATTAGG 60
QY 61 CTGCGCCCGCGCGGCAAGAGTGTCTATGATGAAGCACCCTGGTGGGCGCAGCGCGAG 120
Db 61 TTAAGGCGCAGGGGAAAAAAGCACTATATGATGAACACITAAATATGGGCAAGCAGGAG 120
QY 121 CTGGAGAAAGTTGCGCCCTGAACCCCGCGCTGCTGGAGCAGCAGGAGCTGCAAGCAGATC 180
Db 121 CTGGAAGAAATTGCACTTAACCTTGGCTTTTAGAGACATCAGAAGCTGTAACAAATA 180
QY 181 ATCCGCGCAGCTGCACCCCGCTGCAGACCGGCGCAGCAGGAGCTGAAGAGCCTGTTCAAC 240
Db 181 ATACACAGCTTACACCCAGCTCTCCAGACAGGAACAGAGGAACCTTAGGTCAITATAAT 240
QY 241 ACCGTGCCACCCCTGTACTGCGTGCAGCAGAGAGATCGAGGTCCGCGACACCAAGGAGGCC 300
Db 241 ACAGTAGCAACTCTCTATTGTGTACATGAAAAGATAAAGGTACGAGACACCAAGGAGGCC 300
QY 301 CTGGACAAAGATCGAGGAGGAGCAGAAAGTGCACAGAGATGCCAGAGAGATCCAGCAGGCCAGGCC 360
Db 301 CTAGACAAAGATAGAGGAAGAACAAACAAAGTCAACAAATAACAAAAACAGAAAGCG 360
QY 361 GCCGACAGGGCAAGGTGAGCCAGAACTACCCCATCTGTGCAAGACCTGCAGGGCCAGATG 420

Db 361 ACTGGC---GGAAAGGTCAGTCAAAATATCTATAGTGCAGAAATCTCCAAAGGGCAATG 417
QY 421 GTGCACAGGCCATCAAGCCCGCGCACCCCTGAACGCCCTGGGIGAAGGTGATCGAGGAGAAG 480
Db 418 GTACACAGGCTATATCACCTAGAACCTTTGAATGCATCGGTAAAAGTAATAGAGAGAAAG 477
QY 481 GCCITCAGCCCGGAGGTGATCCCATGTTCCACCGCCCTGAGCGAGGGCGGCACCCGACAG 540
Db 478 GGTTCACCCAGAGGTAAATACCATGTTTACAGCAATATCAGAAGGAGCCACCCACAA 532
QY 541 GACCTGAACACGATGTTGAACACCGTGGCGGGCCACAGGCCCGCCATGCAGATGCTGAAG 600
Db 538 GATCTAAACACCATGTTAATACAGTGGGGGACATCAAGCAGCCCATGCAATGTTAANA 597
QY 601 GACACCATCAACGAGGAGCGCGCGGAGTGGGACCGCGTGCACCCCTGCACGCGCGGCC 660
Db 598 CATACCATCAATGAGGAAGCTGCAGATGGGATAGGTTACATCCAGTACATCCAGGCGCT 657
QY 661 ATCGCCCGCGCCAGATGGCGGAGCCCGCGGACCGGACATGCGCCGACACACAGCAGC 720
Db 658 ATTGCACCGGCCAAATAAGAGAACCAAGGGAAGTGACATAGCAGGAAGTACTGTATCC 717
QY 721 CTGCAGGAGCAGATCGCCTGGATGACCAAGCAACCCCGCATCCCGCTGGCGGACATCTAC 780
Db 718 CTTCAGGAACAATAGCATGCGATGACAAATACCCACCTATTCCAGTGGGAGACATCTAT 777
QY 781 AAGCGGTGGATCATCTCGGCCCTGAACAGATCGTGGGATGTACAGCCCGCGTGAGGATC 840
Db 778 AAAAGATGATANTCTCGCGCTAAATAAATAGTAAGAAATGTACAGCCCTGTCAAGCAAT 837
QY 841 CTGGACATCAAGCAGGCGCTCCCAAGGAGCGCTTCCCGGACTACGTGGAGCGCTTCTTCAG 900
Db 838 CTGGACATAAACAAGGAGCAAGGAACCCCTTAGGGACTAGTAGATCGGTCTTTTAA 897
QY 901 ACCCTGCGCGCGGAGCAGCAGCACCAGGAGGTGAAGAACTGGATGACCGACACACCTGCTG 960
Db 898 ACTTTAAGAGCTGAACAAGGTACACAGATGT-AAAAATGGATGACAGACACACCTTGTG 956
QY 961 GTGCAGAACCGCCAAACCCGACTGCAAGACCATCTCGCGGCTCTCGGCGCCCGGCGCAGC 1020
Db 957 GTTCAAATGCGAACCCAGATGTAGAGCCATTTAAGGCGATTAGGACCGAGGGCTACA 1016
QY 1021 CTGGAGGAGATGATGACCGCTGCCAGGGCGTGGCGGCGCCCGCAGCCACAGGCGCGGCTG 1080
Db 1017 TTAGAAGAATGATGACAGCATGTGAGGAGTGGGGGACCTTGGCCACCAAGCAAGAT 1076
QY 1081 CTGGCGGAGCGATGAGCCAGG---CCACACACAGCGCTGATGTCAGAGAGCAACTTC 1137
Db 1077 TGGCTGAAGCAANTGAGCCAAAGTAACAATAACATAATATGATGCAGAAAGCAATTT 1136
QY 1138 AAGGGCCCGCGCGCATCGTCAAGTGCTTCAACTGGGCAAGGAGGGCGGCACATCGCGCGC 1197
Db 1137 AAAGGCCCTAAAAGAATTGTTAAATGTTTCAACTGIGGCAGGGAAGGGCATATAGCCAGG 1196
QY 1198 AACTGCGCGCGCCCGCGAAGAAGGGCTGCTGGAAGTGGCGCAAGGAGGGCGCCACAGATG 1257
Db 1197 AATTGCAGGGCTCCTGGGAAAAAAGGCTGTGGAATGTGGAAGGAAGGAGACACCAATG 1256
QY 1258 AAGGACTGACCGGAGCGCCAGGCCAACTTCTGGGCAAGATCTGGCCCGACCAAGGCG 1317
Db 1257 AAAGACTGTACTGAGAGACAGGCTAATTTTAGGGAATTTGGCCCTCCCGAGAAGGGG 1316
QY 1318 CGCCCGCGCAACTTCTCTGACAGCGCGCGCGGAGCCCGCCAGCCCGC----- 1359
Db 1317 AGGCGGGGAATCTCTTCAGAACAGACCGAGGCCAACAGCCCGCCACAGCTCCCAACAGCC 1376
QY 1360 CCCCCCGCGGAGAGCTTCCGCTTCGAGAGACACCCCGCGCGCAAGGAGGAGGAGCAAG 1419
Db 1377 CCACGACGAGAGAGCTTCAGGTTGAGGAGACACCCCTGCCCCGAGGCGGAGGAGGANA 1436
QY 1420 GACCGCGAGACCTTGACCAAGCTGAAGACCTGTTCGGCAACGACCCCTGAGCCAGTAA 1479
Db 1419 GTTCTTAAACACCATGTTTAAATACAGTGGGGGAGACATCAAGCAGCCATGCAATGTTTAAA 1473
QY 601 GACACCATCAACGAGGAGCGCGCGGAGTGGGAGTGGGAGTGCACCCCGTGACCGCGCGCC 660

Db 1437 GACAAGGAACCCITAACTGCCCTCAAATCACTCTTTGGACGCGACCCCTTGCTCTCAATAA 1496
RESULT 10
US-09-184-418C-9
; Sequence 9, Application US/09184418C
; Patent No. 6492110
; GENERAL INFORMATION:
; APPLICANT: Hahn, Beatrice
; APPLICANT: Gao, Feng
; APPLICANT: Shaw, George
; TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
; TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS TYPE 1
; FILE REFERENCE: D6287
; CURRENT APPLICATION NUMBER: US/09184,418C
; CURRENT FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 112
; SEQ ID NO 9
; LENGTH: 8972
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; OTHER INFORMATION: isolate=962M751.3; 137.1632:qag; 1419.4435:pol;
; OTHER INFORMATION: 4380.4958:vif; 4898.5188:vpr; 5169.7814:tat;
; OTHER INFORMATION: 5308.7938:rev; 5407.5667:vpu; 5585.8128:env;
; OTHER INFORMATION: 8130.8753:nef
US-09-184-418C-9

Query Match 47.4%; Score 701.2; DB 4; Length 8972;
Best Local Similarity 68.8%; Pred. No. 9.3e-98;
Matches 1032; Conservative 0; Mismatches 443; Indels 25; Gaps 4;
QY 1 ATGGCGCGCGCGCGCGCGATCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
Db 137 AIGGTGCGAGAGCGCTCAATATTAAAGAGCGGCGGAAANTTAGATGAATGGGAAAGATTAGG 196
QY 61 CTGGCGCGCGCGCGCGCGAGGTGCTACATGATGAAGCACCTGGTGTGGCGCGCGCGGAG 120
Db 197 TTAAGCGG 256
QY 121 CTGGAGAAAGTTCGCCCTGACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGATC 180
Db 257 CTGAAAGAAATTTCACCTTACCCCTGGCGCTTTTACAGACATCAGAGGCTGTAACAAATA 316
QY 181 ATCG 240
Db 317 ATACAAAGCTACACACCGCTCTCCAGACAGGAGGAACTTAGTCTATTATAT 376
QY 241 ACCGTGGCG 300
Db 377 ACAGTAGCAACTCTCTATTGTGATGATGAAGAGATAAAGGTACGAGACACCAAGGAGCG 436
QY 301 CTGACAAAGATCGAGGAGGAGGAGCAACAAAGTCCAGCAGAGATCCAGCAGCGCGGAGCG 360
Db 437 CTAGCAAGATAGAGGAGGAGCAACAAAGTCAACAAATAACAAATAACAAAGGAGCG 496
QY 361 GCCGACAAAGGCGCAAGGTGAGCGCAAGCACTACCCCATCGTGCAGAACCTGCAGGCGCAGATG 420
Db 497 ACTGGC---GGAAAGGTCAGTCAAAATTTATCCTATAGTGCAGATCTCCAGGCGCAATG 553
QY 421 GTGCACCG 480
Db 554 GTACACCGCGCTATATCACCTAGAACTTTGAATGCGATGGGTAAAGTAATAGAGGAGAG 613
QY 481 GCCTTCAGCG 540
Db 614 GGTTCACCCAGAGGTAAATACCATGTTTACAGCATTTATCAGAGGAGCGCCACCCACAA 673
QY 541 GACCTGAACACGATGTTGAACACCGTGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
Db 674 GATCTAAACACCATGTTTAAATACAGTGGGGGAGACATCAAGCAGCCATGCAATGTTTAAA 733
QY 601 GACACCATCAACGAGGAGCGCGCGGAGTGGGAGTGGGAGTGCACCCCGTGACCGCGCGCC 660

Db 734 GATACCATCAATGAGGAAGCTGCAAGATGGGATAGTTACATCCAGTACATGCAAGGCGC 793
QY 651 ATGCGCCCGCGGCGAGATGCGCGAGCCCGCGGAGCGACATCGCGGACCAAGCAAGC 720
Db 794 ATTGCACCGGCGCAATCAAGAGAACCCAGGGGAAGTGACATACAGGAAGTACGTAC 853
QY 721 CTGACGAGGACAGATGCGCTGATGACCGCAAGCCCGCCATCGCGTGGGAGACATCTAC 750
Db 854 CTCGAGGAACAATAGCATGATGATGACAAATAACCCACCAATTCACATGGGAGACATCTAT 913
QY 781 AAGCGGTGATATCTGCGCTGGAACAAGATCGTGGGATGATACAGCCCGGTGAGCATC 840
Db 914 AAAAGATGATATCTGCGGTAAATATAAATAGTAAGAAATGATAGATCGGTCTTAAA 973
QY 841 CTGGACATCAAGCAGGCGCCCAAGGAGCCCTTCGCGACTACGAGCCCGGTCTTCAAG 900
Db 974 CTGGACATAAACAAGGCGCAAGGAACCCCTTAGGGAATATGATCGGTCTTAAA 1033
QY 901 ACCCTGCGCGCGGAGAGAGACCCAGGAGGTGGAAGAACTGGATGACCGACACCCCTGCTG 960
Db 1034 ACTTTAAGAGCTGAACAAGCTACACAAGATGT-AAAAATTGGATGACAGACACCTTGTG 1092
QY 961 GTGCAGAACGCCCAACCCGACTGCAAGACCATCTCGCGCTCTCGCCCGCGCGCCAGC 1020
Db 1093 GTTCAAAATGCCAACCAGATTTGTAAGACCATTTTAAGGGCATTAGGACCGCGGCTACA 1152
QY 1021 CTGGAGGAGATGATGACCGCTGCCAGGCGTGGGCGGCCCGCCAGCCACCAAGCGCGG 1080
Db 1153 TTAGAAGAAATGATGACAGCATGTGAGGGAGTGGGGGACCTGSCCAACAAGCAAGATG 1212
QY 1081 CTGGCGGAGGCGATGAGCCAGG---CCAACACAGCGTGTATGATGACAGAAAGCAACTTC 1137
Db 1213 TTGGCTGAAGCAATGAGCCCAAGTAACACATACAAACATTAATGATGCAGAAAGCAATT 1272
QY 1138 AAGGGCCCCCGCGCATCGTCAAGTGCTTCAACTGCGGCAAGGAGGCGCCAGATGCGCCGC 1197
Db 1273 AAAGGCCCTAAAGAAATTTTAAATGTTTCAACTGTCAGGGAAGGCGCATATAGCCAGG 1332
QY 1198 AACTGCGCGCGCGCGCAAGAGGCGTGTGGAAGTGCGGCAAGGAGGCGCCACCAAGATG 1257
Db 1333 AATTGCAAGGCTCTTGGGAAAAAGGCTGTTGGAATGTGGAAGGAAGGACACCCNAATG 1392
QY 1258 AAGGACTGCACCGAGCGCCAGGCAACTTCTCTGGCAAGATCTGGCCCGACCAAGGCG 1317
Db 1393 AAAGACTGTACTGAGAGACAGGCTAAATTTTITAGGGAAATTTGGCCTTCCGAGAAGGG 1452
QY 1318 CGCCCGCGCAACTCTCTGAGAGCGCGCGCGAGCCCAAGCGC----- 1359
Db 1453 AGCCCGGGAACTCTCTCAGAACAGACCCAGAGCCCAACAGCCCGCACAGCTCCACAGCG 1512
QY 1360 CCGCCCGCGGAGAGCTTCCGCTGAGGAGACACCCCGCGGCAAGGAGGAGGAGCAAG 1419
Db 1513 CCACGAGCAGAGAGCTTCAAGTTCGAGGAGACCAACCCCTGCGCGGAGGAGCAGAGAAA 1572
QY 1420 GACCGCGGAGACCTGACAGCGCTGAAGAGCGCTGTCGGCAAGCAACCCCTGAGCCAGTAA 1479
Db 1573 GACAAGGAACCTTAACTGCCCTCAAAATCACCTTTGGCAGGAGCCCTTGTCTCAATAA 1632

RESULT 11

US-09-184-418C-29

; Sequence 29, Application US/09184418C

; Patent No. 6492110

; GENERAL INFORMATION:

; APPLICANT: Hahn, Beatrice

; APPLICANT: Gao, Feng

; APPLICANT: Shaw, George

; TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN

; TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS TYPE 1

; FILE REFERENCE: D6287

; CURRENT APPLICATION NUMBER: US/09/184,418C

; CURRENT FILING DATE: 1999-11-02

; NUMBER OF SEQ ID NOS: 112
; SEQ ID NO 29
; LENGTH: 1503
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; OTHER INFORMATION: isolate=90CR056; gene=gag; CDS:
; OTHER INFORMATION: join(137..1434,1434..4443)
US-09-184-418C-29

Query Match 44.6%; Score 659.8; 33 4; Length 1503;
Best Local Similarity 67.0%; Pred. No. 1.6e-91;
Matches 1007; Conservative 0; Mismatches 472; Indels 24; Gaps 4;

QY 1 ATGGGCGCGCGCGCAGCATCTCTGCGCGGCGCAAGCTGGACSCCTGGGAGCGCATCCGC 60
Db 1 AIGGGTGGGAGAGCGTCACTATTAAAGCGCGGGAATATAGATGCTTGGGAGAAAATTCGG 60
QY 61 CTGGCGCGCGCGCGCAAGTGTCTACATGATGAAGCACCTGGTGTGGGCCAGCCGCGAG 120
Db 61 CTAAGCGCAGGGGGAAGAAAATAATAGGCTAAACATCTAGATGGGCAAGCAGGAG 120
QY 121 CTGGAGAAATTCGCCCTGAACCCCGGCTCTGGAGACCAGGCGGTGCAAGCAGATC 180
Db 121 CTGGAAGATTTGCATTAACCCCGGCTTTTAGAACAACCAAGAGGCTGTCTACAGATA 180
QY 181 ATCCGCCAGCTGCACCCCGCCCTGCAGACCGCGCAGCGAGGAGCTGAAGAGCCTGTCAAC 240
Db 181 ATAGACAGATACAGCCAGCTATTAAAGACAGGAACAGAAACTTAATATTTAAT 240
QY 241 ACCGTGGCCACCCCTGTACTGCTGCAGCAGAAAGATCGAGGTCGCGGACACCAAGAGGCC 300
Db 241 CTAGTAGCAGTCTCTATTGCGTACATCGCAAAATATAGATGTAAAGACACCAAGAGGCT 300
QY 301 CTGGACAAGATCGAGGAGGAGCAGACAAAGTCCACAGAAAGATCCAGCAGGCGGAGGC- 359
Db 301 TTAGATAAGATAGAGGAATAACAAACAAAGTCAGCAAAAACACAGCAAGCAGCAGCT 360
QY 360 -----CGCCGACAAAGGCAAGGTGAGCCAGAAACTACCCCATCTGTGCAGAACCTGCAGGCG 414
Db 361 GATAAGGAAAAAGACAAACAGGTCACTCAAAATTTATCTATAGTACAGAACTGTCAGAGG 420
QY 425 CAGATGTGCACCGGCGCATCAGCCCGCGCACCTCTGACGCGCTGGTGAAGGTGATCGAG 474
Db 421 CAGATGGTACACCGGCGCATATCACCTAGGACCTTAATGATGCATGGGTAAAGTAGTAGAA 480
QY 475 GAGAAGGCTTCAGCCCGGAGGTGATCCCATGTTTCACGCGCCCTGAGCGAGGCGCCACC 534
Db 481 GAAAAGGCTTTIAGCCAGAAAGTAATACCCCATGTTTTCAGCATTAICAGAAAGGAGCCACC 540
QY 535 CCCCAGGACCTGAACACGATGTTGAACACCGTGGCGGCCACCAAGCCCATGCAATG 594
Db 541 CCACAAGACTTAAATGCTATGCTAAATACAGTGGGCGGACATCAAGCAGCCATGCAGATG 600
QY 595 CTGAAGGACACCATCAACGAGGAGGCGCGGAGTGGACCGCTGCACCCCGTGCAGCGCC 654
Db 601 ITAAAGATACAAATCAATGAGGAAGTGCAGAAATGGGACAGGTTACATCCAGTGCATGCA 660
QY 655 GGCCCATGCGCCCGCGCAGATGCGCGAGCCCGCGGAGCGGACATCGCCGCGCACACC 714
Db 651 GGGCCTATTCACCCAGGCCAAATGAGAGAACCAAGGGGAAAGCGATATAGCAGGAACACT 720
QY 715 AGCACCTGCAGGAGAGATCGCCTGATGACCAAGAAACCCCGCATCCCGTGGCGGAC 774
Db 721 AGTACCTGCAGGAACAAATAGCATGGATGACAGGCAATCCAGCTATCCAGTGGGAGAC 780
QY 775 ATCTACAAGCGGTGGATCATCTCTGGCCTGAACAAGATCTGCGGATGTACAGCCCGTG 834
Db 781 ATCTATAAAGATGATGAATCTCTGGGATTAATAAGATAGTAAGAAATGATAGTCTCTGTC 840
QY 835 AGCATCCTGGACATCAAGCAGGCGCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTC 894
Db 841 AGCATTCTGGACATAAAACAAGGGCCAAAAGAACCCCTTTAGAGACTATGTAGACAGGTTT 900

FILE REFERENCE: D6287
CURRENT APPLICATION NUMBER: US/09/184.418C
CURRENT FILING DATE: 1999-11-02
NUMBER OF SEQ ID NOS: 112
SEQ ID NO 48
LENGTH: 4307
TYPE: DNA
ORGANISM: Human immunodeficiency virus type 1
FEATURE:
OTHER INFORMATION: isolate=92N3003; gene=qag-pol fusion
US-09-184-418C-48

Query Match 44.4%; Score 655.6; DB 4; Length 4307;
Best Local Similarity 66.9%; Pred. No. 4.7e-92;
Matches 1005; Conservative 0; Mismatches 474; Indels 24; Gaps 4;

QY 1 ATGGGGCCCGCCGAGCATTCCTGGCGGGGCAAGCTGGACGCTGGGAGCGCATCGGC 60
DB 1 ATGGGTGGAGAGCGTCACTATTAAAGCGGGGGAATAGATGCATGGGAAAAATTCGG 60

QY 61 CTGGCCCCCGGGGCAAGAGTGTCTACATGATGAAGCACCTGGTGGGGCCAGCGCGAG 120
DB 1 TGAGGCCAGGGGAAAGAAAATATAGAAATGAACATTAGTATGGGCAAGCAGGAA 120

QY 121 CTGGAGAGTTCGCCCTGAACCCCGGCTGCTGGAGACCGAGGAGGCTGCAACAGATC 180
DB 121 CTGGAGAGATTCGACCTTAACCTGACCTCTTAGAACAACAGAGGTTGTACGCAATA 180

QY 181 ATCCGCCAGCTGCACCCCGCCCTGCAGACCGGGCAGCGAGGAGCTGGAAGCCTGTCAAC 240
DB 181 ATGACACAGCTGCACCCCTGCTCCAGACAGGAACAGAGGAGATTAAATCATTATTAA 240

QY 241 ACCGTGGCCACCCCTGTACTGGGTGCACGAGAAGATCCAGGTCGGCCACACCAAGAGGCC 300
DB 241 ACAGTAGCAACCCCTCTATTGTACATCAAGAGGATAGAGGTAAACACACCAAGAGCT 300

QY 301 CTGGACAAGATCGAGGAGGAGCAGAACAAAGTCCAGCAGAGATCCAGCAGGC-----C 354
DB 301 CTAGAGGAAGTGGAAAAATACAAAAGAACAGTCAAGAAAGAACACAGAGGCGCAATG 360

QY 355 GAGGCCGCCACAAGGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGC 424
DB 361 GGTAAAGGAACAGCAGCCCAAGTTAGCCAAAATTTATCTATAGTGCAGAAATGCACAAGG 420

QY 415 CAGATGGTGACACAGGCCATCAGCCCGCCACCCCTGAGAGGCTGGAAGTGAAGTGAAG 474
DB 421 CAAGTGGTACACAGCCCATATCACCCTAGGACTTTAAATGCATGGGTAAAGTAAAGAA 480

QY 475 GAGAAGGCTTCAGCCCGGAGGTGATCCCATGTTACCGCCCTGAGCGAGGGCGCACCC 524
DB 481 GAAAGAACTTCAGTCCAGAAATATACCCATGTTTACAGCATTTACAGAGGAGGCCACC 540

QY 535 CCCCAGGAGCTGAACACGATGTTGAACCGGIGGGCGCCACCAAGCGCCCATGAGAGTG 594
DB 541 CCACAAGATTGGAATACCATGCTAAACACCGTGGGGGGCAATCAAGCAGCTATGCAANTG 600

QY 595 CTGAAGGACACCATCAACGAGGAGGCGCGGAGTGGGACCGCTGCACCCCGGTGACCGCC 654
DB 601 CTAAAGATTCTATTATGAAGAGCTGCAGAGTGGGATAGGCTTACATCCACACAGGCCA 660

QY 655 GCGCCCATCGCCCGGCGAGATGGCGAGCCCGCGGAGGAGGAGATCGCGGGACCAACC 714
DB 661 GGACCTATTCCACAGCGCCAGATTAACATGGATGACCAAGGGGAAGTGATATAGCAGGA 720

QY 715 AGCACCTTCAGGAGCAGATCGCCCTGGATGACCAAGCAACCCCGCCATCCCGGTGGGGAC 774
DB 721 ACTACCTTCAGGAAACAATAACATGGATGACCAAGGCCAACCCACCTATCCAGTGGGAG 780

QY 775 ATCTACAAGCGTGGATCATCTCGGCTGGAACAAGATCGTGGGATGTACAGCCCGGTG 834
DB 781 ATTTATAAAGATGGATAATCTCTGGGTTAAATAAATAGTAGAGATGTATAGCCCTGTC 840

QY 835 AGCATCTGGACATCAAGCAGGGGCCCCAAGAGGCCCTTCGCGGACTACGTGGACCCGCTTC 894

DB 841 AGTATTTAGACATAAAACAAGGGCCAAAAGACCTTCAGAGATTATGTGGATAGGTTTC 900

QY 895 TCAAGACCTTGGCGCGGAGGAGCAGACCCAGAGGTGAAGAACTGGATGACCGACACC 954

DB 901 TTTAAACTTTGAGAGCTGAGCAAGCCACACAGAGAGGTAAAAAACTGGATGACAGACACC 960

QY 955 CTGCTGGTGACAGAACCCCAACCCGACTGCAAGACCATCTCGCGGCTCTCGGCCCGGC 1014

DB 961 TTGTTGGTCCAAAATCGGAACCCAGATTGTGAAGACCATCTTAAGAGCATTTAGAGCAGGA 1020

QY 1015 GCCAGCCTGGAGGAGATGATGACCCGCTGCCAGGGCGTGGGGGCTCCCAACCAAGGCC 1074

DB 1021 GCTACACTAGAAGAAATGTTGACACCAIGTCAAGAGAGTGGGAGGACCCAGCCCAAGCA 1080

QY 1075 CGCGTGGTGGCGGAGGAGTGAAGCCAGGCCA-----ACACCAAGCTGATGATGCAG 1125

DB 1081 AGAGTTTATAGCTGAGGCAATGAGCCCGGCAACAGGTACATCAGCAGCCATATGATGCAG 1140

QY 1126 AAGAGCAACTTCAAGGGCCCCCGGCGCATCGTCAAGTCTTCAACTGCGGCAAGGAGGC 1185

DB 1141 AAAAAAATTTAAGGGCCCCGAGACAGGTATTAAGTGTTCACACTGTGGCAAGGAAGGA 1200

QY 1186 CACATGCGCCGCAACTGCGCGCCCCCGGCAAGAGGGCTGCTGAAAGTGGCGCAAGGAG 1245

DB 1201 CATCTAGCCAGAAATTCAGGGGCCCTAGGAAAAAAGGCTGTTGSAATGIGGAAAGGAG 1260

QY 1246 GGCCACAGATGAAGGACTGCACCGAGCGCCAGGCAACTTCCCTGGGCAAGATCTGGCCC 1305

DB 1261 GGACATCAAAATGAAGACTGCACAGAGAGACAGGCTAATTTTIAAGGAAAAATTTGGCCT 1320

QY 1306 AGCCACAAGGGCGCCCGGCAACTTCCCTGCAGAGCGCCCGAGCCCAAGCCCGCCCGCC 1365

DB 1321 TCCACAAGGGGAGGCGCAGGGAATTTCTTTCAGAACAGGCGCAGAGCCCAACAGCCCGCC 1380

QY 1366 GCCGAGAGCTTCCGCTTC-----GAGGAGACCCACCCCGGCGCAGAGGAGGAGCAAGGAC 1422

DB 1381 GCAGAGAGCTTCGGGTTCGGAGAGGAGATAGCCCTTCCCTGAAGCAGGAGCGCGAGGAA 1440

QY 1423 CGCGAGACCC-----TGACCAGCTGAAGAGCCTGTTGCGCAACGACCCCTGAGCCAG 1476

DB 1441 AAGGAATCACCCTCCATTAACTCCCTCAANTCACTCTTTGGCAACGACCCCTAGTCACAG 1500

QY 1477 TAA 1479
DB 1501 TAA 1503

Search completed: September 22, 2003, 22:53:41
Job time : 107.531 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model:

Run on: September 22, 2003, 17:31:08 : Search time 3366.85 Seconds
(without alignments)
10676.533 Million cell updates/sec

Title: US-09-475-704A-3
Perfect score: 1479
Sequence: 1 atggcgccqccqccagcat.....acgacccctgagccagtaa 1479

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562764

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST: *
1: em_estba: *
2: em_esthum: *
3: em_estin: *
4: em_estmu: *
5: em_estov: *
6: em_estpl: *
7: em_estro: *
8: em_htc: *
9: gb_est1: *
10: gb_est2: *
11: gb_htc: *
12: gb_est3: *
13: gb_est4: *
14: gb_est5: *
15: em_estfun: *
16: em_estom: *
17: em_gss_hum: *
18: em_gss_inv: *
19: em_gss_pln: *
20: em_gss_vrt: *
21: em_gss_fun: *
22: em_gss_mam: *
23: em_gss_mus: *
24: em_gss_pro: *
25: em_gss_rod: *
26: em_gss_phg: *
27: em_gss_vrl: *
28: gb_gss1: *
29: gb_gss2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81.2	5.5	749	BQ744212	BQ744212 WHE4113_A
2	78	5.3	807	CB673622	CB673622 OSJNE080
3	78	5.3	821	CB662794	CB662794 OSJNE070
4	78	5.3	824	CB666192	CB666192 OSJNEd12P

5	78	5.3	851	14	CB645622	CB645622 OSJNEb07H
6	76.6	5.2	807	14	CB673355	CB673355 OSJNE07N
7	76.4	5.2	812	14	CB681973	CB681973 OSJNEf08M
8	76.4	5.2	832	14	CB684834	CB684834 OSJNEf14L
9	76.2	5.2	759	10	BF259495	BF259495 HVSMEf0C1
10	75.6	5.1	562	10	BF484304	BF484304 WHE2321_F
11	75.6	5.1	649	14	CB870888	CB870888 HC15J04w
12	75.4	5.1	834	14	CB678700	CB678700 OSJNEf01C
13	75.4	5.1	840	14	CB673814	CB673814 OSJNEe081
14	74.8	5.1	766	14	CB629976	CB629976 OSJNEb06K
15	74.8	5.1	797	14	CB658304	CB658304 OSJNEc14E
16	74.8	5.1	810	14	CB618374	CB618374 OSJIEa02D
17	74.8	5.1	823	14	CB627740	CB627740 OSJIEb020
18	74.8	5.1	826	14	CB641789	CB641789 OSJNEb017
19	74.8	5.1	838	14	CB659169	CB659169 OSJNEc150
20	74.8	5.1	865	14	CB628660	CB628660 OSJIEb04F
21	74.6	5.0	818	14	CB652454	CB652454 OSJNEc02G
22	74.2	5.0	785	10	BG369139	BG369139 HVSMEi002
23	74.2	5.0	933	29	CC391778	CC391778 PUHJ707H
24	73.6	5.0	791	14	CB650382	CB650382 OSJNEb14M
25	73.4	5.0	731	13	BQ752847	BQ752847 WHE4119_G
26	72.6	4.9	2598	11	AY103647	AY103647 Zea mays
27	72.4	4.9	648	10	BE517305	BE517305 WHE0616_A
28	72.2	4.9	803	14	CB633391	CB633391 OSJIEb12D
29	72.2	4.9	1308	11	AY104577	AY104577 Zea mays
30	71.8	4.9	764	13	BQ804819	BQ804819 WHE3559_C
31	71.2	4.8	782	14	CB661708	CB661708 OSJNEd04P
32	71	4.8	787	14	CB658685	CB658685 OSJNEc14O
33	70.6	4.8	2299	11	AY106831	AY106831 Zea mays
34	70.6	4.8	738	13	BU295509	BU295509 603608565
35	70.6	4.8	887	14	CB656981	CB656981 OSJNEc11P
36	70.6	4.8	1146	11	AY104406	AY104406 Zea mays
37	70.4	4.8	784	14	CB650686	CB650686 OSJNEb15F
38	70.4	4.8	792	14	CB618305	CB618305 OSJIEa02B
39	70.2	4.7	794	14	CB658056	CB658056 OSJNEc13N
40	70.2	4.7	2624	11	AY110050	AY110050 Zea mays
41	70	4.7	1491	11	AY104431	AY104431 Zea mays
42	69.8	4.7	637	12	BI960118	BI960118 HVSME002
43	69.8	4.7	925	29	CNS0091P	AL053013 Drosophil
44	69.6	4.7	841	29	CC415062	CC415062 PUHNN18TB
45	69.6	4.7	848	29	CC415063	CC415063 PUHNN18TD

ALIGNMENTS

RESULT 1
BQ744212 749 bp mRNA linear EST 17-JUL-2002
LOCUS WHE4113_A01_A01s Wheat salt-stressed root cDNA library Triticum
DEFINITION aestivum CDNA clone WHE4113_A01_A01, mRNA sequence.
ACCESSION BQ744212
VERSION BQ744212.1 GI:21890999
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
1 (bases 1 to 749)
REFERENCE Anderson,O.D., Akhunov,E., Chao,S., Crossman,C., Deal,K., Dvorak,J.,
AUTHORS Lazo,G.R., Pham,J., Rausch,C.J., Wilson,C. and Woo,J.
TITLE The structure and function of the expressed portion of the wheat
genomes - Salt-stressed root cDNA library
JOURNAL Unpublished
COMMENT Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oanderson@pw.usda.gov
Sequences have been trimmed to remove vector sequence and low

Db 330 CATGGGCTTCGTCCGACGAGCTGGGCTCGGCTCGAGCGGACCCCTGCAAGTGTCTGTCMA 389

QY 504 CATGTTACCGGCTGAGCGAGGGCGCCACCCCGCAGGACTGAACACGATGTTGAACAC 563

Db 390 CATCGAGCAGCAGTCCCGGACATCGCGGAGGGGTGCACGGCCACTTCACCAAGCGGCC 449

QY 564 CQTGGGGCCACACGCGCCGATGAGATGCTGAAGGACACCATCAACGAGGAGGCGGC 523

Db 450 CGAGGAGATCGGCGCGCGGACGAGGCGGACATGTTCCGCTACGCCACCGACGAGACGCC 509

QY 624 CGAGTGGGACCGGCTGCACCCGCTGCACGCGGCGCCCATCGCCCGCGCCGAGATGCGCGA 683

Db 510 CGAGCTGATGCGCCCTCAGCCAGCTCTCGCCACCAAGCTCGCGCGCGCCGCTACCGGAGT 569

QY 684 GCGCGCGGCGAGGACATCGCGGCGCACCCAGCACCCCTCGAGGAGCAGATCGCCCTGGAT 743

Db 570 CCGCAAGAAGCGCACCTCGGCTGGCTCAGGCGCGGACCGGCGGACCGGCTACCGCTTGA 629

QY 744 GACGAGCAACCGCCCATCGCGCTGGGCGGACATCTACAGCGGTGGATCATCTCGGCT 803

Db 630 GTACCTCAACGAGCGCGCGGCGCATGTTCCCGCTCGCGCTCCACACCGCTCATCTCCAC 689

QY 804 GAACAAGATCTGCGGATGTACAGCCCGGCGGAGCATCTGAGCATCAAGTAGGCGCCCA 863

Db 690 CCAGCAGCAGCAGACCGTCAACAACGAGAGATCGCGCGCGCACCTCAAGGAGCAGTCAI 749

QY 864 GGAGCCCTTCGCGGACTAGTGGACCGCTTCTTCAAGACCT 905

Db 750 CAAGCCGCTCATCCCGACAGTACCTCGAGGAGAGACCAT 791

RESULT 3
CB662794
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
CONTACT: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aaa cga cgg cca gtg
BACKWARD: gga aac agc tat gac cat g
Plate: 07 row: C column: 09
Seq primer: gta aaa cga cgg cca gtg.
Location/Qualifiers
1. .821
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNEd07C09"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"

CB662794 821 bp mRNA linear EST 09-APR 2003
OSJNEd07C09.f OSJNEd Oryza sativa (japonica cultivar-group) cDNA
clone OSJNEd07C09 5', mRNA sequence.

CB662794
EST.
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
; (bases 1 to 821)
Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
Unpublished
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aaa cga cgg cca gtg
BACKWARD: gga aac agc tat gac cat g
Plate: 07 row: C column: 09
Seq primer: gta aaa cga cgg cca gtg.
Location/Qualifiers
1. .821
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNEd07C09"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"

/clone_lib="OSJNEd"
/note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
XhoI; 24 hrs after inoculation with Rice Blast (C9240-1)."
BASE COUNT 166 a 309 c 230 g 116 t
ORIGIN

Query Match 5.3%; Score 78; DB 14; Length 821;
Best Local Similarity 44.4%; Pred. No. 0.00074;
Matches 312; Conservative 0; Mismatches 390; Indels 0; Gaps 0;

QY 204 GCAGACCGCGCAGGAGCTGAAGAGCCCTGTTCAACACCGTGGCCACCCCTGACTGCGT 263

Db 88 GCAGATAGAAATGGCGCGGAGACGTTCTCTTACCTCGAGTCCGTGAACGAGGGTCA 147

QY 264 GCAGGAGAAGATCGAGGTCGCGCGACACCAAGAGGCGCCCTGGACAAGATCGAGGAGGACA 323

Db 148 CCGGACAAGCTGTGCGACCAAGGTGTGCGACGCGGTGCTCGACGCGTGCCTCGCCACGA 207

QY 324 GAACAAGTGCCAGCAGAAGATCCAGCAGGCGCGAGGCGCGGACAGGGCAAGGTGAGCCA 383

Db 208 CCGGACAGCAAGGTGGCGTGGAGAGACGTGCACCAAGACCAACATGTTGTTGTCGG 257

QY 384 GAACCTACCCCATCTGTCAGAACCTTCAGGGCCAGATGTTGCACCAAGGCCATCAGCCCCCG 443

Db 268 CGAGATACCAACCAAGGCCACCGTCGACTACGAGAAGATCGTCCGCGACACCTGCGCGCG 327

QY 444 CACCTGAAACGCTGGTGAAGGTGATCGAGGAGAAGGCTTTCAGCCCCGAGGTGATCCC 503

Db 328 CATCGGCTTCGTGTCGACGACGCTGCGGCTCGACGCGCGACCGCTGCAAGGTGCTGCTCAA 387

QY 504 CATGTTACCGCCCTGAGCGAGGCGGCGCCACCCCGCAGGACCTGAACACGATGTTGAACAC 563

Db 388 CATCGAGCAGCAGTCCCGCGACATCGCGCGAGGGGTGCACGCGCCTTCAACCAAGCGGCC 447

QY 564 CGTGGGGCGCCACGAGCGCGCATGTCAGATGTCGAAGGACACCATCAACGAGGAGGCGCG 623

Db 448 CGAGGAGATCGGCGCGCGCGACCGCCACATGTTGCGGTACGCGCACCGACGAGACCCC 507

QY 624 CGAGTGGGACCGGTGACACCCCGTGCACCGCGCGGCGCCCATCGCCCGCGCGAGATGCGCGA 683

Db 508 CGAGCTGATGCCCCCTCAGCCACGCTCTCGCCACCAAGCTCGGCGCGCGCCCTCACCGAGGT 567

QY 684 GCGCGCGCGCAGCAGCATCGCGGCGACCCACCGACCGCTTCAGAGAGCAGATCGGCTGGAT 743

Db 568 CCGCAAGAAGCGCACCTGCGCTGGCTCAGGCGCGGACCGGCAAGACCCAGGTACCGTTGA 627

QY 744 GACGAGCAACCCCGCATCCCGTGGCGGACATCTACAGCGGTGATCATCTGCGGCGCT 803

Db 628 GTACCTCAACGAGCGCGCGCGCATGTTCCCGTCCGCGTCCAGACCGCTCTCACTCCAC 687

QY 804 GAACAAGATCGTGGGATGTACACCCCGTGAAGGATCTTGGACATCAAGCAGGCGCCCAA 863

Db 688 CCAGCAGCAGGAGACCGTCAACAAGCAGGAGATCGCGCGCGCACCTCAAGGAGCAGTCAAT 747

QY 864 GGAGCCCTTCGCGGACTACGTGGACCGCTTCTTCAAGACCT 905

Db 748 CAAGCGGTCATCCCGGACAAAGTACCTCGAGCAGGAGAGACCAT 789

RESULT 4
CB666192
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE

CB666192 824 bp mRNA linear EST 09-APR-2003
OSJNEd12P19.f OSJNEd Oryza sativa (japonica cultivar-group) cDNA
clone OSJNEd12P19 5', mRNA sequence.

CB666192
EST.
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 824)

Db 165 CCGGACAAAGCTGTGGACAGGCTGGACCGGGTGTCTGGACGGGTGCTCGGGCCAGGA 224
QY 324 GAACAGTCCAGCAGCAAGATCCAGCAGGCGCGAGGCGCGGACAGGGCAAGGTGAGCCA 383
Db 225 CCCCAGACGAGGTTGGCTGGAGACGTTGACCAAGACCAACATGTTGTTGG 284
QY 384 GAACATACCCCATCTGTCAGAACTGCAGGGCCAGATGTTGCACAGGCCATCAGGCCCG 443
Db 285 CGAGATCACCACCAAGGCCACCGTCCACTACGAGAGATGTTCCCGACACCTGCCGCG 344
QY 444 CACCCTGAACCGCTGGTGAAGGTGATCGAGGAGAGAGGCTTCAAGCCCGGAGGTGATCCC 503
Db 345 CATCGGCTTCGTGTCGACGACGTCGGCTCGACCGGACCGCTGCAAGGTGCTGCTCAA 404
QY 504 CATGTTACCGCCCTGAGCGAGGGCGCCACCCCGGACCTGACACGATGTTGAACAC 563
Db 405 CATCGAGCAGCAGTCCCGGACATCGCGGACGCGGTCACCGGCACTTCAACAGCGCCC 464
QY 564 CGTGGCGGCCACCGCGCCCATGATGCTGAGGACACCATCAACAGGAGGCGCC 623
Db 465 CGAGGAGATCGCGCGCGCGGACCGAGGCGCCACATGTTGGCTACGCGACCGAGACCCC 524
QY 624 CGAGTGGGACCGCTGACACCCCGTGCAGCGCGCGCCCATCGCCCGGCGGAGATCGGCGA 683
Db 525 CGAGCTGATGCGCCCTCAGCCACGTCCTCGCCACCAAGCTCGGCGCGCGCTCAGCGAGST 584
QY 684 GCGCGCGGCGGAGCATGCGCGGACCGGACCGACCTGCGAGGAGCAGATGCGCTGGAT 743
Db 585 CCGCAAGAACGGCACTGCGCCCTGGCTCAGGCGCGGACCGGACGAGGTCAGCGCTGA 644
QY 744 GACCAAGAACCCCGCCCGTGGCGGACATCTACAGCGGTGGATCATCTGGGCT 803
Db 645 GTACCTCAGCAGCGCGCGCGCATGTTCCCGCTCGCGCTCCACACCGCTCCCTCAGC 704
QY 804 GAACAGATGTTGGGATGTACAGCGCGGTGAGCATCTGAGCATCAAGGAGGCGGCAA 863
Db 705 CGAGCAGCAGGAGACGCTCAGCAAGCAGGAGATGCGCGCGGACCTCAAGGAGCAGCTCAI 764
QY 864 GGAGCGCTTCGGGAGTACGTTGGAGCGGCTTCTTCAAGACCT 905
Db 765 CAAGCGGTTCATCCCGACAACTACTCTGACGAGAGACCAT 806

RESULT 6
CB673355
LOCUS
DEFINITION OSJNE07N10.f OSJNEe Oryza sativa (japonica cultivar-group) cDNA
clone OSJNEe07N10 5', mRNA sequence.
ACCESSION CB673355
VERSION CB673355.1 GI:29677080
KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 807)
AUTHORS Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
TITLE Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
JOURNAL Unpublished
COMMENT Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aac cga cgg cca gta
BACKWARD: gga aac agc tal gac cat g

Plate: 07 row: N column: 10
Seq primer: gta aac cga cgg cca gta
Location/Qualifiers
1..807
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNEe07N10"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNEe"
/note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
XhoI; 24 hrs after inoculation with Rice Blast (70-15)"
BASE COUNT 167 a 301 c 229 g 110 t
ORIGIN

Query Match 5.2%; Score 76.6; DB 14; Length 807;
Best Local Similarity 44.3%; Pred. No. 0.0013;
Matches 310; Conservative 0; Mismatches 389; Indels 0; Gaps 0;
QY 204 GCAGACCGGACGAGGAGCTGAAGACCTGTTTCAACACCGTGGCCACCTGTAAGCT 263
Db 109 GCAGATAGAAATGGCGGAGAGACGTTCTCTTCACTCCGAGTCCGTGAACGAGGCTA 166
QY 264 GCAGGAGAGATCGAGGTCGGGACACCAAGAGGAGGCGCTGGACAGATCGAGGAGGA 323
Db 159 CCGGACACGCTGTGGACACGAGTGTGCGACCGGCTGTCGACCGGTCGCGCCAGGA 228
QY 324 GAACAAGTCCAGCAGAGATCCAGCAGCGGAGGCGCGGACAGGGAAGGTGAGCCA 383
Db 229 CCGGACAGCAAGGTGGCTGCGAGACGTGTCACCAAGACCAACATGGTGTGCTCGG 288
QY 384 GAACATACCCCATCTGTCAGAACCTGCGAGGCGGACAGATGTCACCGAGGCGGCG 443
Db 289 CGAGATCACCACCAAGGCGGACCGCTGACTAGGAGAGATGTCGCGGACACCTGCGCGG 348
QY 444 CACCCTGAAGCGCTGGTGAAGGTGATCGAGGAGAGAGGCTTCAAGCCCGAGGTTGATCC 503
Db 349 CATCGGCTTCGTGTCGACGACGTCGGCTCGACGCGGAGCGGCTGCAAGGTGCTCGTCAA 408
QY 504 CATGTTACCGCGCTGAGCGAGGCGGCGGACCGCGCGGAGGAGTGAACAGATGTTGAACAC 563
Db 409 CATCGAGCAGCAGTGGCGGACATCGCGGAGGCGGCTGCAAGGCGGCTTCAACCAAGCGCC 468
QY 564 CGTGGCGGCGGACCGAGGCGGCGCATGCGAGATGCTGAAGGAGACACCATCAAGGAGGCGCG 623
Db 459 CGAGGAGATCGGCGGCGGCGGACCGGCGGACATGTTGCGCTACGCGGAGGAGGAGCGCC 528
QY 624 CGAGTGGGACCGCTGACCGCGCGTGCAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 683
Db 529 CGAGCTGATGCGCGCTTCAGCGACGCTCTCTCGGACCAAGCTCGGCGGCGGCGGCGGCT 588
QY 684 GCGCGGCGGAGGAGATCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 743
Db 589 CCGCAAGACGCGCACCTGCGGCTGGCTCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 648
QY 744 GACCAAGACCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 803
Db 649 STACCTCAAGCAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 708
QY 804 GAACAAGATCGTGGGATGTACAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 863
Db 709 CCAGCAGCAGGAGACCGCTACCAACGAGAGATCGCGCGGCGGCGGCGGCGGCGGCGG 768
QY 864 GGAGCGCTTCGGGAGTACGTTGGAGCGGCTTCTTCAAGAC 902
Db 769 CAAGCGGTTCATCCCGACAAAGTACCTCGACGAGAGAC 807

RESULT 7
CB681973

```

LOCUS       CB681973                812 bp      mRNA      linear      EST 09-APR-2003
DEFINITION   OSJNEf08M09.f OSJNEf Oryza sativa (japonica cultivar-group) cDNA
clone OSJNEf08M09 5', mRNA sequence.
ACCESSION   CB681973
VERSION     CB681973.1  GI:29685598
KEYWORDS    EST.
SOURCE      Oryza sativa (japonica cultivar-group)
ORGANISM    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE   1 (bases 1 to 812)
AUTHORS     Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
TITLE       Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
JOURNAL     Unpublished
COMMENT     Contact: Rod Wing
            Arizona Genomics Institute
            University of Arizona
            Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
            85721-0088, USA
            Tel: 520 626 3967
            Fax: 520 621 9288
            Email: http://genome.arizona.edu
PCR PRIMERS
FORWARD: gta aaa cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 08 row: M column: 09
Seq primer: gta aaa cga cgg cca gtc.
            Location/Qualifiers
            1..812
            /organism="Oryza sativa (japonica cultivar-group)"
            /mol_type="mRNA"
            /cultivar="Nipponbare"
            /db_xref="taxon:39947"
            /clone="OSJNEf08M09"
            /tissue_type="leaf"
            /dev_stage="3 week"
            /lab_host="DH10B"
            /clone_lib="OSJNEf"
            /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
            XhoI; Uninfected Control"

BASE COUNT   167 a   304 c   228 g   113 t
ORIGIN
Query Match          5.2%; Score 76.4; DB 14; Length 812;
Best local Similarity 44.3%; Pred. No. 0.0014;
Matches 311; Conservative 0; Mismatches 391; Indels 0; Gaps 0;

QY  204 GCACACCGGCGAGGAGGAGCGCTGTAAGAGCGCTGTCAACACCGTGGCCACCGCTGTACAGCT 263
Db  105 GCAGATAGAAATGGCGGCGAGACGCTTCCTCTCACCTCCAGTCCGTGAACGAGGTCA 164

QY  264 GCACGAGAGATCGAGGTCCGCGACACCAAGAGGCGCCCTGGACACAGATCGAGGAGGACCA 323
Db  165 CCGGACAAAGCTGTGGACACAGGTGTCCGACCGGTGTCTGACCGCTGCTGCCCAAGCA 224

QY  324 GAACAAGTCCACGACAGAGATCCAGCAGCGCGAGCGCCGCGACAAAGGCAAGGTGAGCCA 383
Db  225 CCGGACAGCAAGGTGGCGTCCGAGACGTCGACCAAGACCAACATGGTGTGTTGCGG 284

QY  384 GAATACCCCATCGTCAGAACCTGCAGGGCGAGATGGTGCACACAGGCGCATCAGCCCGG 443
Db  285 CGAGATCACCAACCAAGGCCACCGTCGACTACGAGAAGATCGTCCGCGACACCTCCCGCGG 344

QY  444 CACCGTGAACCGCTGGTGAAGTGTATCGAGAGAGAGGCGCTTCAGCCCGGAGGTGATCCC 503
Db  345 CATCGGCTTCGTGTCCGACGACGTCGGCTCGACGCGGACCGCTGCAAGGTGCTGTTCAA 404

QY  504 CATGTTACCGCGCTGAGCGAGCGCGGCCACCGCCAGGACCTGAACACAGATGTGAAAC 563
Db  405 CATCGACGACGACGTGCCCGACATCGCGGAGGGGTGCACGGGCCACTTCACCAAGCGGCC 464

```

```

QY  564 CGTGGCGGCGCACACGAGCGCGCCATCGATGCTCAAGGACACCATCAACGAGGAGCGCGC 623
Db  465 CGAGGAGATCGCGCGCGCGACACGAGGCGCCACATGTTGGCTACGCCACGAGAGACCCC 524

QY  624 CGAGTGGACCGCGTGCACCCCGTGCACCGCGCGCCCATCGCCCGCGCGCCAGATCGCGCA 683
Db  525 CGAGTGTATGCCCTCAGCCACGCTCTGCCACCAAGCTCGCGCGCGCGCTCACCGAGGT 584

QY  684 GCCCGCGCGCAGGACATCGCGCGCACCAACGACACCGCTTCAGGAGGACGATCGCCTGGAT 743
Db  585 CCGCAACAACGGCACTGCGCGCTCAGCGCGCGCGCGCGCAAGACCCAGGTCACCGITGA 644

QY  744 GACCAACAACCCCCCATCCCCGCTGGCGGACATCTACAAAGCGGTGGATCATCTCTGGGCT 803
Db  645 GTACCTCAACGAGCGCGCGCGCGCTGCTCCCGCTCCCGCTCCACACCGCTCTCCATCTCCAC 704

QY  804 GACCAASATCGTCCGATGTACAGCGCGCGTGGACATCTCTTCAAGACCGCT 905
Db  705 CCAGCAGCAGGAGACGCTCACCAACGAGAGATCGCGCGCGGACCTCAAGGAGCAGCTCAT 764

QY  864 GGAGCGCTTCGCGGACTAGTGGACCGCTCTCTTCAAGACCGCT 905
Db  765 CAAGCGGTCTATTCGCGACAAAGTACCTCGACGAGAGAGACCAT 806

```

```

RESULT 8
CB684834
LOCUS      OSJNEf14L24.f OSJNEf Oryza sativa (japonica cultivar-group) cDNA
clone OSJNEf14L24 5', mRNA sequence.
ACCESSION   CB684834
VERSION     CB684834.1  GI:29688559
KEYWORDS    EST.
SOURCE      Oryza sativa (japonica cultivar-group)
ORGANISM    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE   1 (bases 1 to 832)
AUTHORS     Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
TITLE       Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
JOURNAL     Unpublished
COMMENT     Contact: Rod Wing
            Arizona Genomics Institute
            University of Arizona
            Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
            85721-0088, USA
            Tel: 520 626 3967
            Fax: 520 621 9288
            Email: http://genome.arizona.edu
PCR PRIMERS
FORWARD: gta aaa cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 14 row: I. column: 24
Seq primer: gta aaa cga cgg cca gtc.
            Location/Qualifiers
            1..832
            /organism="Oryza sativa (japonica cultivar-group)"
            /mol_type="mRNA"
            /cultivar="Nipponbare"
            /db_xref="taxon:39947"
            /clone="OSJNEf14L24"
            /tissue_type="leaf"
            /dev_stage="3 week"
            /lab_host="DH10B"
            /clone_lib="OSJNEf"
            /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
            XhoI; Uninfected Control"

BASE COUNT   168 a   315 c   231 g   118 t
ORIGIN

```


Query Match		5.2%	Score 76.4;	DB 14;	Length 832;
Best Local Similarity		44.3%	Pred. No. 0.0014;		
Matches 311;		Conservative 0;	Mismatches 391;	Indels 0;	Gaps 0;
QY	204	GCAGACCGGAGGAGGAGCTGAAGAGCGCTGTTCAACACCGTGGCCACCGCTGACTGCGT	263		
Db	97	GCAGATAGAAATGGCGGGAGACGTTCTCTTCACCTCCGAGTCCGIGACGAGGTCA	156		
QY	264	GCACGAGAGNTCGAGGTCCGGCACACCAAGGAGGGCCCTGACAGATCGAGGAGGCA	323		
Db	157	CCCGACACAGCTGTCCGACAGGTTCTCGACCGGTGCTCGACCGTCCCTCGCCACGA	216		
QY	324	GAACAAGTGCCAGCAGAGATCCAGCGCCGAGGCGCCGACACGACAGGAGTACCA	383		
Db	217	CCCCGACAGAGGTGGGTGCGAGACGTGCACCAAGACCAACATGATGATGTTGG	276		
QY	384	GAACATACCCCATCGTGCAGAACCTGACGGGCCAGATGGTGACACCGCCATCAGCCCG	443		
Db	277	CGAGATCACCAACCAAGGACCGTTCGACTACGAGAGATGTCGGGACACCTGCGCCCG	336		
QY	444	CACCTGACCGCCCTGAGCGAGGGCGGACCCGACCCGACCGCTTCAGCGCCGAGTATCC	504		
Db	337	CATCGGCTTCGIGICCGAGGACGTTCGGCTCGACCGGACCGCTGCAAGGTGCTGTCAA	396		
QY	504	CATGTTACCGCCCTGAGCGAGGGCGGACCCGACCCGACCGTGAACAGCATGTTGACAC	563		
Db	397	CATCGAGCAGCATGCGCCGACATCGCGGACGGGGTGCACGGCCACTTCACCAAGCGCC	456		
QY	564	CGTGGCGGCCACACGCGCCCATGACATGCTGAAGACACCATCAGCAGGAGGCGCC	623		
Db	457	CGAGGAGATCGGGCGCGGACGAGGCGGACATGTTGGGCTACGGCACCGAGGAGCGCC	516		
QY	624	CGAGTGGAGCGGCTGACACCGCGGTCACCGCGGCGCCGATCGCCCGCCAGATCGCGCA	683		
Db	517	CGAGCTGATGCCCCCTCAGCCACGTCCTCGCCACCAAGSTCGGGCGCCCTCACCGAGGT	576		
QY	684	GCCCCGGCGAGGACATCGCCGCGACACCGACAGCAGCTGCAGGAGAGATCGCCCTGGAT	743		
Db	577	CGCAAGAGAGCGACCTGCGCGCTGCTAGGCGCGGACAGGAGAGCTGACCGGTGA	636		
QY	744	GACCAAGACCCCGCCATCCCGCTGGGCGACATCTACAGCGGTGGATCATCCGAGCCCT	803		
Db	637	GTACCICACGACGCGCGGCCATGTTGTCGCGTCCGCGTCCACACCCCTCTCATCTCCAC	696		
QY	804	GAACAAGATCGTGGGATGTACAGCCCGCTGAGCATCTCTGGACATCAGCAGGCGCCCAA	863		
Db	697	CCAGCAGCAGAGACCGGTACCAACGACGAGATCGCGCGCGACCTCAGGAGCACATGAT	756		
QY	864	GGAGCCCTTCGGCGACTACGTGGACCGCTTCTTCAGAGCCCT	905		
Db	757	CAAGCCGCTCATCCCGACAGTACCTCGACGAGAGACCAT	798		

RESULT 9
BF259495
LOCUS
DEFINITION
HVSMEf0019H19f Hordeum vulgare seedling root EST library HVCNAC007
(Etiolated and unstressed) Hordeum vulgare subsp. vulgare cDNA
clone HVSMEf0019H19f, mRNA sequence.

ACCESSION
BF259495
VERSION
BF259495.2 GI:13120022
KEYWORDS
EST
SOURCE
Hordeum vulgare subsp. vulgare
ORGANISM
Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Hordeum.
REFERENCE
1 (bases 1 to 759)
Wing, R., Close, T.J., Klein, H., A., Wise, R., Bequim, D., Frisch, D., Yu
, Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Choi, D.W., Fenton
, R.D., Oates, R. and Main, D.
Development of a genetically and physically anchored EST resource

TITLE

JOURNAL
COMMENT
for barley genomics: Morex unstressed seedling root cDNA library
Unpublished
On Nov 16, 2000 this sequence version replaced gi:11188608.
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Total hq bases = 388
Seq primer: AATTAACCTCACTAAAGGG
High quality sequence stop: 758.
Location/Qualifiers
1..759
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Morex"
/db_xref="taxon:112509"
/clone="HVSMEf0019H19f"
/tissue_type="Seedling root"
/lab_host="TJCL12"
/clone_lib="Hordeum vulgare seedling root EST library
HVCNAC007 (Etiolated and unstressed)."
/note="Vector: lambdaZAP; Site_1: EcoRI; Site_2: XhoI.
Seeds were surface sterilized then germinated under axenic
conditions in the dark at room temperature on filter paper
with water, nystatin and cefotaxime in covered
crystallization dishes. Five-day old seedling roots were
then harvested, total RNA was prepared, poly(A) RNA was
purified, one primary unamplified cDNA library was made,
and 1 million pfu were in vivo excised to give pBluescript
SK(-) cDNA phagemids. These steps were performed in the
Close laboratory at the University of California,
Riverside (Choi, Close, Fenton). Phagemids were plated and
picked at the Clemson University Genomics Institute (CUGI)
(Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA
preparations, DNA sequencing and sequence analysis were
performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates
, Rambo, Main). The sequence has been trimmed to remove
vector sequence and contains a minimum of 100 bases of
phred value 20 or above. For more details on library
preparation and sequence analysis see
http://www.genome.clemson.edu/projects/barley. To order
this clone see http://www.genome.clemson.edu/orders/Also
see Close TJ, Wing R, Klein, H., A., Wise R (2001)
Genetically and physically anchored EST resources for
barley genomics. Barley Genetics Newsletter 31:29-30.
(http://wheat.pw.usda.gov/ggpages/bqnl/31/cover.html)"

BASE COUNT 158 a 282 c 184 g 133 t 2 others
ORIGIN

Query Match:		5.2%	Score 76.2;	DB 10;	Length 759;
Best Local Similarity		47.7%	Pred. No. 0.0015;		
Matches 284;		Conservative 0;	Mismatches 305;	Indels 6;	Gaps 2;
QY	124	GAGAAGTTCGCCCTGAACCCCGCTGCTGGAGACACCGAGGCTGCAAGCAGATCATC	183		
Db	13	GAGATCACCGCAACCAACCATGGAACCTGATGACGACGTCGGCGGACGAGGCCAGTTC	72		
QY	184	CGCCAGCTGCACCCCGCCCTGCAGACCGGCGAGGAGGCTGAAGAGCCTGTCAACACC	243		
Db	73	CTCAACATGCTGCTCAAGCTCATCGGCGCAAGAGACCATGGAGATCGGCGCTACACG	132		
QY	244	GTGGCCACCCCTGACTGCTGGTCACGAGAGATCGAGGTCCGCGACACCAAGGAGCCCTG	303		
Db	133	GGTACTCCCTCICGCCACCGCTCGCCCATCCCCGACGACGCGCACCATCTTGCCCATG	192		
QY	304	GACAAGATCGAGGAGGAGCAGAACAGTGCCACGAGAGATCCAGCAGGCGGCGCCGCC	363		
Db	193	GACATCAACCGCGAGAGACTACGAGCTGGGGCTGCCGTGTCATCGAGAGGCGCGGTGGCG	252		
QY	364	GACAAGGCAAGGTGAGCCAGAACTACCCCATCTGTCAGACACTGCGAGGCGCAGATGGTG	423		

[illegible]

RESULT	10
BF484304	
LOCUS	BF484304
DEFINITION	WHE2321_F12_K23%S wheat pre-anthesis spike cDNA library Triticum aestivum cDNA clone WHE2321_F12_K23, mRNA sequence.
	linear EST C6-BEC-2000
	562 bp

BASE COUNT		108 a	215 c	154 g	85 t	<p>preparations and DNA sequencing were performed in the OI Anderson lab (all other authors).</p>			
ORIGIN						Score 75.6;	DB 10;	Length 562;	
Query Match						Best Local Similarity 50.4%;	Pred. No. 0.0018;		
Matches 270;		Conservative	0;	Mismatches 254;	Indels	12;	Gaps	33	
QY	182	TCGCGCAGCTGCACCCCGCCCTGCAGACCGGCAGCGAGGASCTGAAGAGCCCTGTTCAACA	241						
LB	26	TCCTCAACATGCTGCTCAAGCTCATCGCGGCCAAGAAGACCATGGAGATCGCGTCTACA	85						
QY	242	CGGTGCGCACCCCTGTACTGCGTGCCACGAGAGATCGAGGTCCGCGGACACCAAGGAGGCC	301						
DB	85	CGGGTACTCCCTGCTCGCCACCGCGCTCGGCATCCCGAGCAGCGGCACCATCTTGGCCA	145						
QY	302	TGGACAGATCGAGGAGGAGCACAACAAGTSCCAGCAGAGATCCAGCAGGCCGAGGCCG	361						
DB	146	TGGACATCAACCCGAGAACIACTATCIGGGGCTCCCGTGCATCGAGAGGCCGCGGTGG	205						
QY	362	CCGACAAGGGCAAGGTGAGCCACAACACTACCCCATCGTGCAGAACCTGCAGGGGCCAGATGG	421						
DB	206	CGCACAGATCGAGCTTCGCGGAGG---CCCGCGCTGCGCGCTCTGGACGCGCTGCTGG	262						
QY	422	TGCACCAGGCCATCAGCCCCCGCACCCCTGAACGCGCTGGGTAAAGTGATCGAGGAGAGG	481						
DB	263	AGGACGAGGCCAACACCGGACCTTCGACTTCGTCTCTGTTGGACCGCAGCAAGGACAACT	322						
QY	482	CCTTCAGCCCCGAGTGATCCCATGTTACGCGCCCTGAGCGAGGCGCCACCCCGCAGG	541						
DB	323	ACCTCAACTACCAGAGCGGCTCATGAAGCTGTCGAAGGTGGCGGCCCTCTCTCGGCTACG	382						
QY	542	ACCTGAACACGATGTGAACACCGTGGGGGGCCACCAGGCCGCCCATGCAGATGCTGAAGG	601						
DB	383	AC---AACACCCCTTGGACGGGTCCGTCTGCTCCCGCGCGACGCCCCCATGCGCAGT	439						
QY	602	ACACCATCAACGAGGAGGCCCGCCGAGTGGGACCGCGTGCACCCCCGIGCAGCGCCGCCCA	661						
DB	440	ACATCGGCTACIA-----CCGGGACTTCGTCTCGAGCTCAACAAGGCCCTCGCGCGCG	493						
QY	662	TGCCCCCGGCGCAGATCGCGGAGCCCGCGGCGCAGCGACATCGCCGGCACCAACGAG	717						
DB	494	ACCAGCGGTCGAGATCTGCCAGTCTCCCGTCCCGCGCAGCGGCAATCACCCCTCTGCCG	549						

RESULT 11
CH870888

505

DEFINITION

5-PRIME, mRNA sequence.

ACCESSION CB870888

VERSION

KEYWORDS

308105

ORGANISM

Eukaryota; Viridiplantae: Streptophyta

Spermatophyta; Magnoliophyta;

: *Triticeae*: *Hordeum*:

REFERENCE

REFERENCES

varshney, s. k.; zhang, n.; barton, k.
Graner, A.

TITLE Barley ESTs from coleoptile tissue

JOURNAL OF
BARLEY & LENTIL
CROSSLING

JOURNAL
CUPUBLISHED
CONTACT: Stein Nils

COMMENT

CONTACT: STEPHEN NILES

Molecular Markers Group

Department

Genbank

Molecular Markers Group, Department GenBank
Institute of Plant Genetics and Crop Plant Research (IPK)

Institute of Plant Genetics and Crop Plant
Correspondence: 3 06466 Catanzaro, Germany

CORRENSR. 3, 0
-91: 036487-5522Tel: 039482-5522
Fax: 039482-5595

Fax: 039482-5595

Email: stein@ipk-gatersleben.de

Insert Length: 649 Std Err

Db	388	CATCGAGCAGCAUTCGCGGAGACATCGCGGAGGGGTGCACGGCCACTTCACCAAGGAGCC	447
QY	564	CGTGGCGGCGACAGGCGCGCATGCAGATGCTGAAGGACACCATCAACGAGGAGGCGGC	623
Db	448	CGAGGAGATCGGCGCGCGAGCAGGGCCACATGTCGGGTACGCCACCGACGAGAGCCCG	507
QY	624	CGAGTGGACCGGCTGCACCGGTGCACGGCGGCGCCCATCGCCCGCGCCAGATCGGAGA	693
Db	508	CGAGGTGATGCCCGTCAGCGACGTCTCGCCACCAAGGTTCGGCGCGCGCTCACCGAGCT	567
QY	684	GCCCCCGGCGAGGACATCGCGGCGACCAACGACACCTTCGACGAGGAGATCGGCTGAT	743
Db	568	CCGCAAGAACGGCACTCGGCGCTGCGTCAAGGCCCCGACCGCAAGACCCAGGTCAACGTGA	627
QY	744	GACCAACCAACCCGCCATCCCGCTGGCGGACATCIACAGCGGTGGATCACTCTGGCGCT	803
Db	628	GTACCTCAACGACGCGCGCGCATGGTCCCGCTCCGCGTCCACACCGTCTCATCTCCAC	687
QY	804	GAACAGATCGTGGGATGTACAGCCCCGTGAGCATCTGTGACATCAAGACGAGGCGCCCAA	863
Db	688	CCAGCAGCAGGAGACCGGTCACCAACGACGAGATCGCGCGGACCTCAAGGAGCACGTCAI	747
QY	864	GGAGCCCTTCGCGGACTACGTGACCGGCTTCTTCAAGACGCT	905
Db	748	CAAGCGGGTCAATCCCCCGGACAGTACCTCGACGAGAGACCAT	789

Search completed: September 22, 2003, 22:50:10
Job time : 3375.85 secs


```

PR 01-SEP-1999; 99US-0152195.
XX (CHIR ) CHIRON CORP.
XX Barnett S, Zur Megede J;
XX P1
XX WPI; 2000-452401/39.
DR P-PSDB; AAY96943.
XX
XX Polynucleotide encoding antigenic type C HIV Gag polypeptide of a HIV
XX Env polypeptide and the polypeptide useful for immunizing a mammal
XX especially human against HIV
XX
XX Claim 2: Page 92-93; 113pp; English.
XX
XX Expression cassettes comprising a polynucleotide encoding antigenic
XX type C human immunodeficiency virus (HIV) Gag or Env polypeptides are
XX useful in DNA immunization, generation of packaging cell lines and
XX production of Gag- and/or Env-containing proteins. Synthetic Env and Gag
XX expression cassettes exhibit increased potency for induction of
XX cytotoxic T-lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1
XX self-assemble into non-infectious virus-like particles which are used as
XX a matrix for the proper presentation of an antigen entrapped or
XX associated to the immune system of the host.
XX
XX Sequence 1479 BP; 325 A; 529 C; 463 G; 162 T; 0 other;
XX
XX Query Match 100.0%; Score 1479; DB 21; Length 1479;
XX Best Local Similarity 100.0%; Pred. No. 1.6e-186;
XX Matches 1479; Conservative 0; Mismatches 0; Indels 0; Gaps 0
XX
QY 1 ATGGGCGCCCGCGCCAGCATCTCTCGCGCGCGGAAGCTGGACGCTGGAGCGCATCCG 60
DB 1 ATGGCGCCCGCGCCAGCATCTCTCGCGCGCGGAAGCTGGACGCTGGAGCGCATCCG 60
QY 61 CTGCGCCCGCGCGCGGAAGTGTCTACATGATGAAGCACCTGGTGTGGCCACGCCGAG 120
DB 61 CTGCGCCCGCGCGCGGAAGTGTCTACATGATGAAGCACCTGGTGTGGCCACGCCGAG 120
QY 121 CTGGAGAAGTTTCGCCCTGAACCCCGCCCTCTCTGGAGACCAGCGAGGCTGCAAGCAGATC 180
DB 121 CTGGAGAAGTTTCGCCCTGAACCCCGCCCTCTCTGGAGACCAGCGAGGCTGCAAGCAGATC 180
QY 181 ATCCGCGAGCTGCACCCCGCCCTCTCAGACCGGACGAGGAGCTGAAGAGCCTGTTCAAC 240
DB 181 ATCCGCGAGCTGCACCCCGCCCTCTCAGACCGGACGAGGAGCTGAAGAGCCTGTTCAAC 240
QY 241 ACCGTGGCCACCTGTACTCGGTGCACGAGAGATCGAGTCCCGCGACACCAAGAGCGC 300
DB 241 ACCGTGGCCACCTGTACTCGGTGCACGAGAGATCGAGTCCCGCGACACCAAGAGCGC 300
QY 301 CTGGACACAGATCGAGGAGGAGGACAGACAACTACCCCATCTCTCAGAACCTGCAGGCGCCAGATG 420
DB 301 CTGGACACAGATCGAGGAGGAGGACAGACAACTACCCCATCTCTCAGAACCTGCAGGCGCCAGATG 420
QY 361 GCGGACAGGGCAAGTGCAGCCAGAACTACCCCATCTCTCAGAACCTGCAGGCGCCAGATG 420
DB 361 GCGGACAGGGCAAGTGCAGCCAGAACTACCCCATCTCTCAGAACCTGCAGGCGCCAGATG 420
QY 421 GTGCACAGGCCATCAGCCCGCCCGACCCCTGAAGCCCTGGTGAAGGTGATCCAGAGAA 480
DB 421 GTGCACAGGCCATCAGCCCGCCCGACCCCTGAAGCCCTGGTGAAGGTGATCCAGAGAA 480
QY 481 GCCTTCAGCCCGGAGTGTATCCCATGTTTACCGCCCTTGAGCGAGGGCGCCACCCCGAG 540
DB 481 GCCTTCAGCCCGGAGTGTATCCCATGTTTACCGCCCTTGAGCGAGGGCGCCACCCCGAG 540
QY 541 GACCTGAACACGATGTTGAACACCGTGGCGGCCACCAAGCGCCCATGAGATGCTGAAG 600
DB 541 GACCTGAACACGATGTTGAACACCGTGGCGGCCACCAAGCGCCCATGAGATGCTGAAG 600
QY 601 GACACCATCAACGAGGAGGCGCCCGAGTGGGACCGCGTGCACCCCGTGGACCGCGCGCC 660
DB 601 GACACCATCAACGAGGAGGCGCCCGAGTGGGACCGCGTGCACCCCGTGGACCGCGCGCC 660

```

Df		601	GACACCATCAACGAGGAGCGCGCCGAGTGGACCGCGCTGCACCCCGTSCACCCCGGTSCACGCCGCGCCCC	660
Qy		661	ATCGCCCCCGCCAGATGCGCGAGCCCCCGGGCAGCGACATCGCGCGCACCAACAGCACC	720
Dd		661	ATCGCCCCCGCCAGATGCGCGAGCCCCCGGGCAGCGACATCGCGCGCACCAACAGCACC	720
Qy		721	CTGAGGAGCAGATCGCCTGGATGACCAACCCCCCACATCCCCGTGGGCGACATCTAC	780
Dd		721	CTGAGGAGCAGATCGCCTGGATGACCAACCCCCCACATCCCCGTGGGCGACATCTAC	780
Qy		781	AAGCGTGATCATCTTGGCCTGAACAAGATCGTGGGATGTACAGCCCGTGAGCATC	840
Dd		781	AAGCGTGATCATCTTGGCCTGAACAAGATCGTGGGATGTACAGCCCGTGAGCATC	840
Qy		841	CTGSACATCAAGCAGGCGCCCAAGGAGCCCTTCGCGACTAGTGGACCGGTTCTTCAAG	900
Dd		841	CTGSACATCAAGCAGGCGCCCAAGGAGCCCTTCGCGACTAGTGGACCGGTTCTTCAAG	900
Qy		901	ACCCTGCGCGGAGCAGACACCCAGGAGCTGAAGAATCGATGACCGACACCCCTGCTG	960
Dd		901	ACCCTGCGCGGAGCAGACACCCAGGAGCTGAAGAATCGATGACCGACACCCCTGCTG	960
Qy		961	GTGCAGAACGCCAACCOCGACTGCAAGACCATCTGTGGCGCTCTCGGCCCTCGGCCCGC	1020
Dd		961	GTGCAGAACGCCAACCOCGACTGCAAGACCATCTGTGGCGCTCTCGGCCCTCGGCCCGC	1020
Qy		1021	CTSGAGGAGATGATGACCGCTGCCAGGGCTGGCGGCCACAGCCACAGGCCCGCGTG	1080
Dd		1021	CTSGAGGAGATGATGACCGCTGCCAGGGCTGGCGGCCACAGCCACAGGCCCGCGTG	1080
Qy		1081	CTGCCCGAGGCGATGAGCCAGGCCCAACACAGCGCTGATGTCAGAGAGCAACTTCAAG	1140
Dd		1081	CTGCCCGAGGCGATGAGCCAGGCCCAACACAGCGCTGATGTCAGAGAGCAACTTCAAG	1140
Qy		1141	GGCCCCCGGCGCATGTCAGTGTCTCAACTGCGGCAAGGAGGGCCACATCGCCCCAAC	1200
Dd		1141	GGCCCCCGGCGCATGTCAGTGTCTCAACTGCGGCAAGGAGGGCCACATCGCCCCAAC	1200
Qy		1201	TGCCCGCCCCCGCAAGAGGGCTGTGGAGTGGCGCAGGAGGGCCACAGATGAAG	1260
Dd		1201	TGCCCGCCCCCGCAAGAGGGCTGTGGAGTGGCGCAGGAGGGCCACAGATGAAG	1260
Qy		1261	GACTGCACCGAGCGCCAGGCCAACTTCTTGGCAAGATCTGTGGCCAGCCACAGGGCCGC	1320
Dd		1261	GACTGCACCGAGCGCCAGGCCAACTTCTTGGCAAGATCTGTGGCCAGCCACAGGGCCGC	1320
Qy		1321	CCCGGCAACTTCTTGCAGAGCGCCCGGAGGCCACCGCCCCCCCCCGCGGAGAGCTTCGGC	1380
Dd		1321	CCCGGCAACTTCTTGCAGAGCGCCCGGAGGCCACCGCCCCCCCCCGCGGAGAGCTTCGGC	1380
Qy		1381	TTCGAGGAGACCAACCCCGCGCAGAGCAGGAGAGCAAGGACCGCGGAGCCCTGACCAGC	1440
Dd		1381	TTCGAGGAGACCAACCCCGCGCAGAGCAGGAGAGCAAGGACCGCGGAGCCCTGACCAGC	1440
Qy		1441	CTGAAGAGCCCTTCTGGCAAGACCCCTGTGAGCCAGTAA	1479
Dd		1441	CTGAAGAGCCCTTCTGGCAAGACCCCTGTGAGCCAGTAA	1479

RESULT 2
AA-44548
ID- AAL44548 standard; DNA; 1479 BP.
XX
AC AAL44548;
XX
DT 06-NOV-2002 (first entry)
XX
DE HIV-1 p55gag polypeptide coding sequence 1.
XX
KW HIV; ds; vacciner; gene; immune response; microparticle;
KW adsorbent surface; poly(alpha-hydroxy acid); polyhydroxy butyric acid;
KW polycaprolactone; polyorthoester; polycyanacrylate; detergent;
KW submicron emulsion; viral infection; bacterial infection;

KW parasitic infection; HIV-1 p55gag polypeptide.
XX Human immunodeficiency virus type 1.
OS WO200226209-A2.
PN 04-APR-2002.
XX 28-SEP-2001; 2001WO-US30540.
XX 28-SEP-2000; 2000US-236105P.
PR 30-AUG-2001; 2001US-315905P.
XX (CHIR) CHIRON CORP.
PA O'hagan D, Otten G, Donnelly JJ, Polo JM, Barnett S, Singh M;
PI Ulmer J, Dubensky JW;
XX WPI; 2002-519084/55.
DR
XX A microparticle to which a biologically active macromolecule is
PT adsorbed, for use as a vaccine composition to treat viral, bacterial or
PT parasitic infections, comprises a polymer microparticle, a detergent
PT and a submicron emulsion.
XX Claim 72; Fig 1; 100pp; English.
PS
XX The invention relates to a method of raising an immune response in a host
CC animal. The method of the invention comprises administering a
CC microparticle that has an adsorbent surface to which a first biologically
CC active macromolecule (e.g. a nucleic acid) has been adsorbed. The
CC microparticle comprises a polymer microparticle of poly(alpha-hydroxy
CC acid), a polyhydroxy butyric acid, a polycaprolactone, a polyorthoester,
CC a polycyanoacrylate, a detergent, and submicron emulsion. The method/
CC microparticle of the invention is useful for immunising a host animal
CC against viral, bacterial or parasitic infections. The present DNA
CC sequence encodes a HIV-1 p55gag polypeptide.
XX
SQ Sequence 1479 BP; 325 A; 529 C; 463 G; 162 T; 0 other;

Query Match 100.0%; Score 1479; DB 24; Length 1479;
Best Local Similarity 100.0%; Pred. No. 1.6e-136;
Matches 1479; Conservative 0; Mismatches 0; Index 0; Gaps 0;

QY 1 ATGGGCGCGCGCCAGCATCCTCGCGCGCGGCAAGCTGGACGCGCTGGAGCGCATCCCG 60
DB 1 ATGGGCGCGCGCGCGCGCATCCTCGCGCGCGGCAAGCTGGACGCGCTGGAGCGCATCCCG 60

QY 61 CTGCGCGCGCGCGCGCAAGAGTGTACATGATGAAGCACTGGTGGTGGCGCGAG 120
DB 61 CTGCGCGCGCGCGCGCAAGAGTGTACATGATGAAGCACTGGTGGTGGCGCGAG 120

QY 121 CTGGAGAGTTCGCCCTGACCCCGCGCTGAGACGAGCGAGGCTGCAAGCAGATC 180
DB 121 CTGGAGAGTTCGCCCTGACCCCGCGCTGAGACGAGCGAGGCTGCAAGCAGATC 180

QY 181 ATCCGCCAGTGCACCCCGCGCTGAGACGCGAGGAGCTGAAGAGCCTGTTCAAC 240
DB 181 ATCCGCCAGTGCACCCCGCGCTGAGACGCGAGGAGCTGAAGAGCCTGTTCAAC 240

QY 241 ACCGTGGCCAGCCTGTACTGGTGCACGAGAAGATCGAGGTCCGCGACACCAAGGAGCC 300
DB 241 ACCGTGGCCAGCCTGTACTGGTGCACGAGAAGATCGAGGTCCGCGACACCAAGGAGCC 300

QY 301 CTGGACAAGATCGAGGAGGAGCAGAACAACTGCCAGCAGAAGATCCAGAGGCGGAGCC 360
DB 301 CTGGACAAGATCGAGGAGGAGCAGAACAACTGCCAGCAGAAGATCCAGAGGCGGAGCC 360

QY 361 GCGGACAAGGCAAGGTGAGCGCAGAACTACCCCATCGTGCAGAACCTGAGGGCCAGATG 420
DB 361 GCGGACAAGGCAAGGTGAGCGCAGAACTACCCCATCGTGCAGAACCTGAGGGCCAGATG 420

QY 421 GTGCACACAGGCCATCAGCCCGCGCACCCCTGACAGCGCTGGGTGAAGGTGATCGAGGAGAAG 480

DB 421 GTGCACACAGGCCATCAGCCCGCGCACCCCTGAAGCGCTGGGTGAAGGTGATCGAGGAGAAG 480
QY 481 GCCTTCAGCCCGGAGGTGATCCCGCATGTTCACCGCCTGAGCGAGGGCGCCACCCCGCAG 540
DB 481 GCCTTCAGCCCGGAGGTGATCCCGCATGTTCACCGCCTGAGCGAGGGCGCCACCCCGCAG 540
QY 541 GACCTGACACAGGATGTTGAACACCTGTTGGGCGGCCACAGGCGCGCATGCAGATGCTGAAG 600
DB 541 GACCTGACACAGGATGTTGAACACCTGTTGGGCGGCCACAGGCGCGCATGCAGATGCTGAAG 600
QY 601 GACACCATCAACGAGGAGCGCGCGAGTGGGACCGCGTGCACCCCGTGCACCGCGGCGCC 660
DB 601 GACACCATCAACGAGGAGCGCGCGAGTGGGACCGCGTGCACCCCGTGCACCGCGGCGCC 660
QY 661 ATCGCCCGCGCGAGTCCGCGAGCCCGCGCGAGGACATCGCGCGCACATCGCGCGCACACACC 720
DB 661 ATCGCCCGCGCGAGTCCGCGAGCCCGCGCGAGGACATCGCGCGCACATCGCGCGCACACACC 720
QY 721 CTGCAGGACGAGATCGCGTGGATGACGACGACACCCCGCGCATCCCGTGGGGGACATCTAC 780
DB 721 CTGCAGGACGAGATCGCGTGGATGACGACGACACCCCGCGCATCCCGTGGGGGACATCTAC 780
QY 781 AAGCGTGGATCATCCTGGCGCTGAACAAGATCGTGGGATGTACAGCCCGCGTGCAGCATC 840
DB 781 AAGCGTGGATCATCCTGGCGCTGAACAAGATCGTGGGATGTACAGCCCGCGTGCAGCATC 840
QY 841 CTGGACATCAAGCAGGCGCGCGAGGAGCGCTTCGCGACTACGTGGACCGCTTCTTCAAG 900
DB 841 CTGGACATCAAGCAGGCGCGCGAGGAGCGCTTCGCGACTACGTGGACCGCTTCTTCAAG 900
QY 901 ACCCTGCGCGCGGAGCAGACACCGAGAGGTGAAGAACTGGATGACCGGACACCGCTGCTG 960
DB 901 ACCCTGCGCGCGGAGCAGACACCGAGAGGTGAAGAACTGGATGACCGGACACCGCTGCTG 960
QY 961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCTCTCGCGCGCGCGGCGAGC 1020
DB 961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCTCTCGCGCGCGCGGCGAGC 1020
QY 1021 CTGGAGGAGATGATGACCGCTGCGAGGCGGTGGGCGCGCGCGAGCGCGCGCGGTG 1080
DB 1021 CTGGAGGAGATGATGACCGCTGCGAGGCGGTGGGCGCGCGCGAGCGCGCGCGGTG 1080
QY 1081 CTGGCGGAGGCGGATGAGCGAGGCGCGCGAGCGCGGTGATGATGATGATGATGATGATGATG 1140
DB 1081 CTGGCGGAGGCGGATGAGCGAGGCGCGCGAGCGCGGTGATGATGATGATGATGATGATGATG 1140
QY 1141 GCGCGCGCGCGCGCATGCTCAAGTGTCTTCACTGCGGCAAGGAGGCGCGCATGCGCGCGAAC 1200
DB 1141 GCGCGCGCGCGCGCATGCTCAAGTGTCTTCACTGCGGCAAGGAGGCGCGCATGCGCGCGAAC 1200
QY 1201 TCGCGCGCGCGCGCGCAAGAGGCGTGTGGAAGTGGGCAAGGAGGCGCGCGCATGAGTGAAG 1260
DB 1201 TCGCGCGCGCGCGCGCAAGAGGCGTGTGGAAGTGGGCAAGGAGGCGCGCGCATGAGTGAAG 1260
QY 1261 GACTGCACCGAGCGCGCGCGCAACTTCTTGGGCAAGATGCTGGCGCGCGCGCGCGCGCGCG 1320
DB 1261 GACTGCACCGAGCGCGCGCGCAACTTCTTGGGCAAGATGCTGGCGCGCGCGCGCGCGCGCG 1320
QY 1321 CCGCGCAACTTCTTGCAGAGCG 1380
DB 1321 CCGCGCAACTTCTTGCAGAGCG 1380
QY 1381 TTCGAGGAGACCG 1440
DB 1381 TTCGAGGAGACCG 1440
QY 1441 CTGAAGAGCGCTGTTGCGCAACGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1479
DB 1441 CTGAAGAGCGCTGTTGCGCAACGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1479

QY 1381 TCGAGGAGACACCCCGCCAGAACGAGGAGGACCGAGAGCCCTGACCAGC 1449
|||||
Db 1382 TCGAGGAGACACCCCGCCAGAACGAGGAGGACCGAGAGCCCTGACCAGC 1449
|||||
QY 1441 CTGAAGAGCCTGTTCCGCAACGACCCCTGAGCCAGTAA 1479
|||||
Db 1441 CTGAAGAGCCTGTTCCGCAACGACCCCTGAGCCAGTAA 1479
|||||
RESULT 4
ACA03523
ID ACA03523 standard: DNA: 4419 BP.
XX
AC ACA03523;
XX
DT 22-MAY-2003 (first entry)
XX
DE Synthetic DNA encoding immunogenic HIV peptide #6.
XX
KW Immunogenic HIV polypeptide; human immunodeficiency virus; HIV;
KW vaccine; gene therapy; packaging cell line; humoral immune response;
KW cellular immune response; gene delivery vector; DNA immunisation;
KW ds.
XX
OS Synthetic.
XX
PN WO2003004657-A1.
XX
PD 16-JAN-2003.
XX
PF 05-JUL-2002; 2002WO-US21422.
XX
PR 05-JUL-2001; 2001US-303192P.
PR 31-AUG-2001; 2001US-316860P.
PR 16-JAN-2002; 2002US-349728P.
PR 16-JAN-2002; 2002US-349793P.
PR 16-JAN-2002; 2002US-349871P.
XX
PA (CHIR) CHIRON CORP.
XX
PI Zur Megede J, Barnett SW, Lian Y;
XX
DR WPI; 2003-221602/21.
XX
PT New synthetic polynucleotides encoding antigenic HIV type B and/or type
PT C polypeptides, useful as immunogenic compositions or vaccines for
PT generating humoral or cellular immune responses against HIV in a
PT subject, especially humans -
XX
PS Example 1; Fig 11; 262pp; English.
XX
CC The invention describes a synthetic polynucleotide encoding 2 or more
CC immunogenic HIV polypeptides, where at least 2 of the polypeptides are
CC derived from different HIV subtypes. The polynucleotide is useful for
CC immunisation, generation of packaging cell lines, or production of HIV
CC polypeptides. The polynucleotide and its encoded proteins are useful as
CC immunogenic compositions or vaccines for generating humoral or cellular
CC immune responses against HIV in a subject, or for inducing neutralising
CC antibodies against HIV. The gene delivery vector comprising the
CC polynucleotide is also useful for DNA immunisation of, or for
CC generating an immune response (e.g. a humoral or cellular immune
CC response) in, a subject such as a mammal, particularly a human. This
CC sequence encodes a human immunodeficiency virus immunogenic peptide.
XX
SQ Sequence 4419 BP; 980 A; 1582 C; 1371 G; 486 T; 0 other;

Query Match 99.9%; Score 1477.4; DB 25; Length 4419;
Best Local Similarity 99.9%; Pred. No. 2.1e-186;
Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGGCCCGCCGCGGAGCATCCTGCGCGGCGGCAAGCTGGACGCTGGGAGCGCATCCGC 60
|||||

Db 7 ATGGGGCCCGCCGCGGAGCATCCTGCGCGGCGGCAAGCTGGACGCTGGGAGCGCAICCGC 66
QY 61 CTGCGCCCGCGGCGGCAAGAGTGTCTACATGATGAAGCACTGGTGGTGGCCAGCCGCGAG 120
|||||
Db 67 CTGCGCCCGCGGCGGCAAGAGTGTCTACATGATGAAGCACTGGTGGTGGCCAGCCGCGAG 126
|||||
QY 121 CTGGAGAGTTCGCCCTGAACCCCGGCTGCTGGAGACCGAGGCTGCAAGCAGATC 180
|||||
Db 127 CTGGAGAGTTCGCCCTGAACCCCGGCTGCTGGAGACCGAGGCTGCAAGCAGATC 186
|||||
QY 181 ATCCGCCAGCTGCACCCCGCCTGCAGACCGCGCAGCGAGAGCTGAAGAGCCCTGTTCAC 240
|||||
Db 187 ATCCGCCAGCTGCACCCCGCCTGCAGACCGCGCAGCGAGAGCTGAAGAGCCCTGTTCAC 246
|||||
QY 241 ACCGTGGCCACCCCTGTACTGCGTGCACGAGAAGATCGAGTCCGCGACCAAGGAGGCC 300
|||||
Db 247 ACCGTGGCCACCCCTGTACTGCGTGCACGAGAAGATCGAGTCCGCGACCAAGGAGGCC 306
|||||
QY 301 CTGGACAGATCGAGGAGGAGCAGAACAGTCCAGCAGCAAGATCCAGCAGGCCGAGGCC 360
|||||
Db 307 CTGGACAGATCGAGGAGGAGCAGAACAGTCCAGCAGCAAGATCCAGCAGGCCGAGGCC 366
|||||
QY 361 GCCGACAAAGGGAAGGTGAGCCAGAACTACCCCATCTGTCAGAACCTTCAGGGCCAGATG 420
|||||
Db 367 GCCGACAAAGGGAAGGTGAGCCAGAACTACCCCATCTGTCAGAACCTTCAGGGCCAGATG 426
|||||
QY 421 GTGCACCCAGGCCATCAGCCCGCCGACCCCTGAACGCTGGTGAAGGTGATCGAGGAGAAG 480
|||||
Db 427 GTGCACCCAGGCCATCAGCCCGCCGACCCCTGAACGCTGGTGAAGGTGATCGAGGAGAAG 486
|||||
QY 481 GCCTTCAGCCCGGAGGTGATCCCATCTGTCACCGCCCTGAGCGAGGCCGCCACCCCGAG 540
|||||
Db 487 GCCTTCAGCCCGGAGGTGATCCCATCTGTCACCGCCCTGAGCGAGGCCGCCACCCCGAG 546
|||||
QY 541 GACCTGAACACAGATGTTGAACACCGTGGCGGGCCACACCGCCGCTGAGATGCTGAAG 600
|||||
Db 547 GACCTGAACACAGATGTTGAACACCGTGGCGGGCCGCCACCGCCGCTGAGATGCTGAAG 606
|||||
QY 601 GACACCATCAACGAGGAGGCCCGCGAGTGGGACCGCGTGCACCCCTGCACCCCGGCCCC 660
|||||
Db 607 GACACCATCAACGAGGAGGCCCGCGAGTGGGACCGCGTGCACCCCGGCCCGGCCCC 666
|||||
QY 661 ATCGCCCGCGGCCAGATGCGCGAGCCCGCGGAGCGGAGATCGCGGACCGCCACCGACCC 720
|||||
Db 667 ATCGCCCGCGGCCAGATGCGCGAGCCCGCGGAGCGGAGATCGCGGACCGCCACCGACCC 726
|||||
QY 721 CTGCAGGACAGATGCGCTGGATGACCGAGCAACCCCGCCATCCCGTGGCGGACATCTAC 780
|||||
Db 727 CTGCAGGACAGATGCGCTGGATGACCGAGCAACCCCGCCATCCCGTGGCGGACATCTAC 786
|||||
QY 781 AAGCGGTGATCATCTCTGGCCCTGAACAGATCGTCCGGATGTACAGCCCGCGGAGATC 840
|||||
Db 787 AAGCGGTGATCATCTCTGGCCCTGAACAGATCGTCCGGATGTACAGCCCGCGGAGATC 846
|||||
QY 841 CTGGACATCAAGCAGGCGGCCCAAGGAGCCCTTCGCGGACTACGTGGACCGCTTCCTCAAG 900
|||||
Db 847 CTGGACATCAAGCAGGCGGCCCAAGGAGCCCTTCGCGGACTACGTGGACCGCTTCCTCAAG 906
|||||
QY 901 ACCCTGCGGCGCGGAGCAGCACCACCGAGGAGGTGAAGAACTGGATGACCGACACCCCTGCTG 960
|||||
Db 907 ACCCTGCGGCGCGGAGCAGCACCACCGAGGAGGTGAAGAACTGGATGACCGACACCCCTGCTG 966
|||||
QY 961 GTGCAGAACGCCAACCCCGACTGCAGACCACTCTCTGGCGGCTCTCGGCCCGCGGCCAGC 1020
|||||
Db 967 GTGCAGAACGCCAACCCCGACTGCAGACCACTCTCTGGCGGCTCTCGGCCCGCGGCCAGC 1026
|||||
QY 1021 CTGGAGGAGATGATGACCGCCTGCCAGGCGGTGGCGGCCCGCCAGCCACAGGCCCGCGTG 1080
|||||
Db 1027 CTGGAGGAGATGATGACCGCCTGCCAGGCGGTGGCGGCCCGCCAGCCACAGGCCCGCGTG 1086
|||||
QY 1081 CTGGCCGAGGCGATGAGCCAGGCCAACACCGAGGTGATGATGCAGAGAGCAACTTCAAG 1140
|||||
Db 1087 CTGGCCGAGGCGATGAGCCAGGCCAACACCGAGGTGATGATGCAGAGAGCAACTTCAAG 1146
|||||

QY 1141 GGCCCCCGCGCATCGTCAAGTGTCTTCAACTGCGGCAAGGAGGCGCCACATCGCCCGCAAC 1200
Db 1147 GGCCCCCGCGCATCGTCAAGTGTCTTCAACTGCGGCAAGGAGGCGCCACATCGCCCGCAAC 1206
QY 1201 TGGCGCGCCCGCGCAAGAGGGCTGCTGAAGTGGCGCAAGGAGGCGCCACATGATGAAG 1260
Db 1207 TGGCGCGCCCGCGCAAGAGGGCTGCTGAAGTGGCGCAAGGAGGCGCCACATGATGAAG 1266
QY 1261 GACTGCACCGAGCGCGCGCAAGTTCCTGGGCAAGATCTGGCGCGCGCGCAAGGAGGCG 1320
Db 1267 GACTGCACCGAGCGCGCGCAAGTTCCTGGGCAAGATCTGGCGCGCGCGCAAGGAGGCG 1326
QY 1321 CCGGCAACTTCTTGCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
Db 1327 CCGGCAACTTCTTGCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1386
QY 1381 TTCGAGGAGACACCG 1440
Db 1387 TTCGAGGAGACACCG 1446
QY 1441 CTGAAGAGCGCTTTCGGCAACGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1479
Db 1447 CTGAAGAGCGCTTTCGGCAACGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1485

RESULT 5

ACA03522
ID ACA03522 standard; DNA: 4716 BP.

AC ACA03522;

XX 22-MAY-2003 (first entry)

DE Synthetic DNA encoding immunogenic HIV peptide #5.

XX Immunogenic HIV polypeptide; human immunodeficiency virus; HIV;
KW vaccine; gene therapy; packaging cell line; humoral immune response;
KW cellular immune response; gene delivery vector; DNA immunisation;
KW ds.

XX Synthetic.

XX WO2003004657-A1.

XX 16-JAN-2003.

XX 05-JUL-2002; 2002WO-US21421.

XX 05-JUL-2001; 2001US-303192P.

XX 31-AUG-2001; 2001US-316860P.

XX 16-JAN-2002; 2002US-349728P.

XX 16-JAN-2002; 2002US-349793P.

XX 16-JAN-2002; 2002US-349871P.

XX (CHIR) CHIRON CORP.

XX Zur Megede J, Barnett SW, Lian Y;

XX WPI; 2003-221602/21.

XX New synthetic polynucleotides encoding antigenic HIV type B and/or type
PT C polypeptides, useful as immunogenic compositions or vaccines for
PT generating humoral or cellular immune responses against HIV in a
PT subject, especially humans -

XX Example 1: Fig 10; 262pp; English.

XX The invention describes a synthetic polynucleotide encoding 2 or more
CC immunogenic HIV polypeptides, where at least 2 of the polypeptides are
CC derived from different HIV subtypes. The polynucleotide is useful for
CC immunisation, generation of packaging cell lines, or production of HIV
CC polypeptides. The polynucleotide and its encoded proteins are useful as

CC immunogenic compositions or vaccines for generating humoral or cellular
CC immune responses against HIV in a subject, or for inducing neutralising
CC antibodies against HIV. The gene delivery vector comprising the
CC polynucleotide is also useful for DNA immunisation of, or for
CC generating an immune response (e.g. a humoral or cellular immune
CC response) in, a subject such as a mammal, particularly a human. This
CC sequence encodes a human immunodeficiency virus immunogenic peptide.
XX

SQ Sequence 4716 BP; 1041 A; 1684 C; 1463 G; 528 T; 0 other;

Query Match 99.9%; Score 1477.4; DB 25; Length 4716;
Best Local Similarity 99.9%; Pred. No. 2.1e-186;
Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGGCGCCCGCCAGCATCTCTGCGGGCGGCAAGCTGTGACGCTGGAGCGCATCCGC 60
Db 7 ATGGGGCGCCCGCCAGCATCTCTGCGGGCGGCAAGCTGTGACGCTGGAGCGCATCCGC 66
QY 61 CTGGCGCCCGCGGCGCAAGAGTGTACATCATGAAGCACTGGTGTGGCGCAGCGCGAG 120
Db 67 CTGGCGCCCGCGGCGCAAGAGTGTACATCATGAAGCACTGGTGTGGCGCAGCGCGAG 126
QY 121 CTGGAGAGTTGGCCCTGAACCCCGGCTCTGGAGACGAGGGCTGCAAGCAGATC 180
Db 127 CTGGAGAGTTGGCCCTGAACCCCGGCTCTGGAGACGAGGGCTGCAAGCAGATC 186
QY 181 ATCCGCCAGCTGCACCCCGGCTTCAGACCGGCGCAGCGAGAGCTGAACAGCCCTGTCAAC 240
Db 187 ATCCGCCAGCTGCACCCCGGCTTCAGACCGGCGCAGCGAGAGCTGAACAGCCCTGTCAAC 246
QY 241 ACCGTGGCCACCCCTGTACTGCTGCACGAGAAGATCGAGTCCGCGACACCAAGGAGGCC 300
Db 247 ACCGTGGCCACCCCTGTACTGCTGCACGAGAAGATCGAGTCCGCGACACCAAGGAGGCC 306
QY 301 CTGGACAAGATCGAGGAGGAGCAGACAAGTCCAGCAAGATCCAGCAGCGCGAGGCC 360
Db 307 CTGGACAAGATCGAGGAGGAGCAGACAAGTCCAGCAAGATCCAGCAGCGCGAGGCC 366
QY 361 GCGGACAAGGCGAAGGTGAGCCAGAACTACCCCATCTGTCAGAACTCGCAGAGCCAGATG 420
Db 367 GCGGACAAGGCGAAGGTGAGCCAGAACTACCCCATCTGTCAGAACTCGCAGAGCCAGATG 426
QY 421 GTGCACCAAGGCGCATCAGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
Db 427 GTGCACCAAGGCGCATCAGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 486
QY 481 GCCTTCAGCCCGGAGGTGATCCCATGTTACCGCCCTGAGCGAGGCGCCACCCCGCAG 540
Db 487 GCCTTCAGCCCGGAGGTGATCCCATGTTACCGCCCTGAGCGAGGCGCCACCCCGCAG 546
QY 541 GACCTGAACACGATGTTGAACACCGTGGCGGCGCCACCGCGCGCGCGCGCGCGCGCG 600
Db 547 GACCTGAACACGATGTTGAACACCGTGGCGGCGCCACCGCGCGCGCGCGCGCGCGCG 606
QY 601 GACACCATCAACGAGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
Db 607 GACACCATCAACGAGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 666
QY 661 ATCGCCCG 720
Db 667 ATCGCCCG 726
QY 721 CTGCAGGACGATGCGCTGGATGACCAAGCAACCCCGCGCGCGCGCGCGCGCGCGCG 780
Db 727 CTGCAGGACGATGCGCTGGATGACCAAGCAACCCCGCGCGCGCGCGCGCGCGCGCG 786
QY 781 AAGCGGTGGATCATCTGGGCGCTGAACAAGATCGTGGGATGTACAGCCCGCGTACCAATC 840
Db 787 AAGCGGTGGATCATCTGGGCGCTGAACAAGATCGTGGGATGTACAGCCCGCGTACCAATC 846
QY 841 CTGGACATCAAGCAGGCG 900
Db 847 CTGGACATCAAGCAGGCG 906

```
QY 901 ACCCTGCGCGCCGAGCAGACAGACACCCAGAGGIGAAAGAACTGGATGACCGACACACCCCTGGTG 960
Db 907 ACCCTGCGCGCCGAGCAGACACCCAGAGGIGAAAGAACTGGATGACCGACACACCCCTGGTG 966
QY 961 GTGCAGAACCCCAACCCCGACTGCAAGACCATCTGCGCGCTCTCGCGCCCGCGCGCCAGC 1020
Db 967 GTGCAGAACCCCAACCCCGACTGCAAGACCATCTCTGCGCGCTCTCGCGCCCGCGCGCCAGC 1026
QY 1021 CTGGAGGAGATGATGACCGCTGCGCAGCGCTGGCGCGCCCGCCAGACACAGGCGCCGCTG 1080
Db 1027 CTGGAGGAGATGATGACCGCTGCGCAGCGCTGGCGCGCCCGCCAGACACAGGCGCCGCTG 1086
QY 1081 CTGGCGGAGCGGATGAGCGAGCGCCCAACACCCAGCGTGTATGATGATGATGATGATGATGATG 1140
Db 1087 CTGGCGGAGCGGATGAGCGAGCGCCCAACACCCAGCGTGTATGATGATGATGATGATGATGATG 1146
QY 1141 GGCCCCCGCGGCATCGTCAAGTCTTCAACTGCTGCGGAGGAGGCGCCACATCGCGCCCAAC 1200
Db 1147 GGCCCCCGCGGCATCGTCAAGTCTTCAACTGCTGCGGAGGAGGCGCCACATCGCGCCCAAC 1206
QY 1201 TGCCCGCGCGCGCGCAAGAGGGCTGCTGGAAGTGGCGCAAGGAGGCGCCACCATGGAAG 1260
Db 1207 TGCCCGCGCGCGCGCAAGAGGGCTGCTGGAAGTGGCGCAAGGAGGCGCCACCATGGAAG 1266
QY 1261 GACTGACCGAGCGCGCGCGCAACTTCTCTGGGCAAGATCTGGCGCGCGCGCGCGCGCGCG 1320
Db 1267 GACTGACCGAGCGCGCGCGCAACTTCTCTGGGCAAGATCTGGCGCGCGCGCGCGCGCGCG 1326
QY 1321 CCGGCAACTCTCTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
Db 1327 CCGGCAACTCTCTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1386
QY 1381 TTCGAGGAGACACCCCGCGCGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440
Db 1387 TTCGAGGAGACACCCCGCGCGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1446
QY 1441 CTGAAGAGCGTGTTCGGCAACGACCCCTCTGAGCGCAAG 1479
Db 1447 CTGAAGAGCGTGTTCGGCAACGACCCCTCTGAGCGCAAG 1485
```

```
RESULT 6
ACA03524
ID ACA03524 standard; DNA: 2742 BP.
```

```
AC ACA03524;
```

```
XX 22-MAY-2003 (first entry)
```

```
DE Synthetic DNA encoding immunogenic HIV peptide #7.
```

```
XX Immunogenic HIV polypeptide; human immunodeficiency virus; HIV;
```

```
KW vaccine; gene therapy; packaging cell line; humoral immune response;
```

```
KW cellular immune response; gene delivery vector; DNA immunisation;
```

```
KW ds.
```

```
OS Synthetic.
```

```
XX WO200304657-A1.
```

```
PN 16-JAN-2003.
```

```
XX 05-JUL-2002; 2002WO-US21421.
```

```
PR 05-JUL-2001; 2001US-303192P.
```

```
PR 31-AUG-2001; 2001US-316860P.
```

```
PR 16-JAN-2002; 2002US-349728P.
```

```
PR 16-JAN-2002; 2002US-349793P.
```

```
PR 16-JAN-2002; 2002US-349871P.
```

```
XX (CHIR ) CHIRON CORP.
```

```
PA
```

```
XX
```

```
FI Zur Megede J, Barnett SW, Lian Y;
XX WPI: 2003-221602/21.
DR
XX New synthetic polynucleotides encoding antigenic HIV type B and/or type
P: C polypeptides, useful as immunogenic compositions or vaccines for
PT generating humoral or cellular immune responses against HIV in a
PR subject, especially humans -
XX
```

```
FS Example 1; Fig 12: 262pp; English.
```

```
XX The invention describes a synthetic polynucleotide encoding 2 or more
CC immunogenic HIV polypeptides, where at least 2 of the polypeptides are
CC derived from different HIV subtypes. The polynucleotide is useful for
CC immunisation, generation of packaging cell lines, or production of HIV
CC polypeptides. The polynucleotide and its encoded proteins are useful as
CC immunogenic compositions or vaccines for generating humoral or cellular
CC immune responses against HIV in a subject, or for inducing neutralising
CC antibodies against HIV. The gene delivery vector comprising the
CC polynucleotide is also useful for DNA immunisation of, or for
CC generating an immune response (e.g. a humoral or cellular immune
CC response) in, a subject such as a mammal, particularly a human. This
CC sequence encodes a human immunodeficiency virus immunogenic peptide.
```

```
XX Sequence 2742 BP; 578 A; 1020 C; 859 G; 285 T; 0 other;
```

```
Query Match 99.8%; Score 1475.8; DB 25; Length 2742;
```

```
Rest Local Similarity 99.8%; Pred. No. 3.7e-186;
```

```
Matches 1477; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 ATGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
Db 7 ATGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 66
QY 61 CTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Db 67 CTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 126
QY 121 CTGGAGAGTTCGCGCTGAAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
Db 127 CTGGAGAGTTCGCGCTGAAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 186
QY 181 ATCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
Db 187 ATCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 246
QY 241 ACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
Db 247 ACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 306
QY 301 CTGGACAAGATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
Db 307 CTGGACAAGATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 366
QY 361 GCGGACAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
Db 367 GCGGACAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 426
QY 421 GTGCACCGAGGCGCATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
Db 427 GTGCACCGAGGCGCATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 486
QY 481 GCGTTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
Db 487 GCGTTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 546
QY 541 GACCTGAACACGATGTTGACACCGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
Db 547 GACCTGAACACGATGTTGACACCGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 606
QY 601 GACACCATCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
Db 607 GACACCATCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 666
```

```
QY 661 ATCGCCCGCGCCAGATCGCGGAGCCCGCGGCGGCGAGCATCGCCGCGCACCCAGCACC 720
DB 667 ATCGCCCGCGCCAGATCGCGGAGCCCGCGGCGGCGGCGAGCATCGCCGCGCACCCAGCACC 726
QY 721 CTGCAGGACGATCGCCCTGGATGACGACGACACCCCGCCATCCCGCTGGGGGACATCTAC 780
DB 727 CTGCAGGACGATCGCCCTGGATGACGACGACACCCCGCCATCCCGCTGGGGGACATCTAC 786
QY 781 AAGCGGTGAATCATCTCTGCGCTGAACAGATCGTGGGATGATGACGACCCCGTGAGCAAC 840
DB 787 AAGCGGTGAATCATCTCTGCGCTGAACAGATCGTGGGATGATGACGACCCCGTGAGCAAC 846
QY 841 CTGGACATCAAGCAGGCGCTCAAGAGGCGCTTCCGCGACTACGTTGACCGCTTCTTCAAG 900
DB 847 CTGGACATCAAGCAGGCGCTCAAGAGGCGCTTCCGCGACTACGTTGACCGCTTCTTCAAG 906
QY 901 ACCCTGCGCGCGGAGCAGACGACCCAGGAGGTGAAGAACCTGGATGACGACACCCGCTG 960
DB 907 ACCCTGCGCGCGGAGCAGACGACCCAGGAGGTGAAGAACCTGGATGACGACACCCGCTG 966
QY 961 GTGCAGAACGCCAACCCGACTGCAAGACCATCTCGCGGCTCTCGCGCCCGGCGGAGC 1020
DB 967 GTGCAGAACGCCAACCCGACTGCAAGACCATCTCGCGGCTCTCGCGCCCGGCGGAGC 1026
QY 1021 CTGGAGGAGATGATGACCGGCTGCGAGGCGGTGGCGGCGCCGCGACACAGGCGCGGCTG 1080
DB 1027 CTGGAGGAGATGATGACCGGCTGCGAGGCGGTGGCGGCGCCGCGACACAGGCGCGGCTG 1086
QY 1081 CTGGCGAGCGGATGAGCGAGCGCCAACACGAGCGGTGATGCGAGAGCAACTTCAAG 1140
DB 1087 CTGGCGAGCGGATGAGCGAGCGCCAACACGAGCGGTGATGCGAGAGCAACTTCAAG 1146
QY 1141 GCGCCCGCGGAGCATCTGCAAGTCTTCAACTGCGGCAAGGAGGCGGCGACATCGCGCGAAC 1200
DB 1147 GCGCCCGCGGAGCATCTGCAAGTCTTCAACTGCGGCAAGGAGGCGGCGACATCGCGCGAAC 1206
QY 1201 TCGCGCGCGCGCGGAGGCTGCTGGAAGTTCGCGCAAGGAGGCGGCGGCGGAGGAG 1260
DB 1207 TCGCGCGCGCGCGGAGGCTGCTGGAAGTTCGCGCAAGGAGGCGGCGGCGGAGGAG 1266
QY 1261 GACTGCACCGAGCGCGGAGCGCAACTTCTGCGGCAAGATCTGCGCGCGCGGCGGCGG 1320
DB 1267 GACTGCACCGAGCGCGGAGCGCAACTTCTGCGGCAAGATCTGCGCGCGCGGCGGCGG 1326
QY 1321 CCGCGCAACTCTCTGAGAGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1380
DB 1327 CCGCGCAACTCTCTGAGAGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1386
QY 1381 TTCGAGGAGACCGCGCGGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440
DB 1387 TTCGAGGAGACCGCGCGGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1446
QY 1441 CTGAAGAGCGCTGTTGCGGCAAGGACCCCTGAGCGGAGTAA 1479
DB 1447 CTGAAGAGCGCTGTTGCGGCAAGGACCCCTGAGCGGAGTAA 1485
```

```
RESULT 7
ACA03521
ID ACA03521 standard: DNA; 5145 BP.
XX
AC ACA03521:
XX
DT 22-MAY-2003 (first entry)
XX
DE Synthetic DNA coding immunogenic HIV peptide #4.
XX
KW Immunogenic HIV polypeptide; human immunodeficiency virus; HIV;
KW vaccine; gene therapy; packaging cell line; humoral immune response;
KW cellular immune response; gene delivery vector; DNA immunisation;
KW ds.
XX
```

```
OS Synthetic.
XX WC2003004657-A1.
PN
XX 16-JAN-2003.
PD
XX 05-JUL-2002; 2003WO-US21421.
PF
XX 05-JUL-2001; 2001US-303192P.
PR 31-AUG-2001; 2001US-316860P.
PR 16-JAN-2002; 2002US-349728P.
PR 16-JAN-2002; 2002US-349793P.
PR 16-JAN-2002; 2002US-349871P.
XX (CHIR ) CHIRON CORP.
XX Zur Megede J, Barnett SW, Lian Y;
XX WPI; 2003-221602/21.
XX
XX New synthetic polynucleotides encoding antigenic HIV type B and/or type
XX C polypeptides, useful as immunogenic compositions or vaccines for
XX generating humoral or cellular immune responses against HIV in a
XX subject, especially humans -
XX
XX Example 1; Fig 9; 262pp; English.
XX
XX The invention describes a synthetic polynucleotide encoding 2 or more
XX immunogenic HIV polypeptides, where at least 2 of the polypeptides are
XX derived from different HIV subtypes. The polynucleotide is useful for
XX immunisation, generation of packaging cell lines, or production of HIV
XX polypeptides. The polynucleotide and its encoded proteins are useful as
XX immunogenic compositions or vaccines for generating humoral or cellular
XX immune responses against HIV in a subject, or for inducing neutralising
XX antibodies against HIV. The gene delivery vector comprising the
XX polynucleotide is also useful for DNA immunisation of, or for
XX generating an immune response (e.g. a humoral or cellular immune
XX response) in, a subject such as a mammal, particularly a human. This
XX sequence encodes a human immunodeficiency virus immunogenic peptide.
XX
XX Sequence 5145 BP; 1130 A; 1842 C; 1599 G; 574 T; 0 other;
```

```
Query Match 99.8%; Score 1475.8; DB 25; Length 5145;
Best Local Similarity 99.9%; Pred. NO. 3.4e-186;
Matches 1477; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 ATGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
DB 7 ATGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 56
QY 61 CTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 120
DB 67 CTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 126
QY 121 CTGGAGAGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 180
DB 127 CTGGAGAGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 186
QY 181 ATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 240
DB 187 ATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 246
QY 241 ACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
DB 247 ACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 306
QY 301 CTGGACAGATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
DB 307 CTGGACAGATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 366
QY 361 GCCGACAGGCGCAAGGTGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420
DB 367 GCCGACAGGCGCAAGGTGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 426
```

QY	421	GTGCACAGGCCATCAGCCCCCGGACGCTGAACGCGTGGGGAAGGATGATCAGAGAAAG	430
DB			
DB	427	GTGCACAGGCCATCAGCCCCCGGACGCTGAACGCGTGGGGAAGGATGATCAGAGAAAG	486
QY	481	GCCTTCAGCCCGAGGTATCCCCATGTTTACCGGCGTGTAGCGAGGGCGCCACCGCCAG	540
DB			
DB	487	GCCTTCAGCCCGAGGTATCCCCATGTTTACCGGCGTGTAGCGAGGGCGCCACCGCCAG	546
QY	541	GACCTGAACACGATGTTGAACACCGTGGGCGGCCACAGGCGGCCCATGTCAGATGCTGAAG	600
DB			
DB	547	GACCTGAACACGATGTTGAACACCGTGGGCGGCCACAGGCGGCCCATGTCAGATGCTGAAG	606
QY	601	GACACCATCAACGAGGAGCGCGGAGTGGGACCGGTGCACCCGTGCACGCGCGGCC	660
DB			
DB	607	GACACCATCAACGAGGAGCGCGGAGTGGGACCGGTGCACCCGTGCACGCGCGGCC	666
QY	661	ATCGCCCGCGGCAGATGCGCGAGCCCGGCGGACCGGTGCACCCGTGCACGCGCGGCC	720
DB			
DB	667	ATCGCCCGCGGCAGATGCGCGAGCCCGGCGGACCGGTGCACCCGTGCACGCGCGGCC	726
QY	721	CTGCAGGAGCAGATCGCTGGATGACCAAGCAACCCCGCATCCCGTGGGGGACATCTATC	780
DB			
DB	727	CTGCAGGAGCAGATCGCTGGATGACCAAGCAACCCCGCATCCCGTGGGGGACATCTATC	786
QY	781	AAGCGGTGGATCATCCTGGGCCGAGCAAGATCGTGGGATGTACAGCCCGGAGGACATC	840
DB			
DB	787	AAGCGGTGGATCATCCTGGGCCGAGCAAGATCGTGGGATGTACAGCCCGGAGGACATC	846
QY	841	CTGGACATCAAGCAGGCGGCCAAGGAGCCCTTCGCGACTACCTGGACCGCTCTCTCAAG	900
DB			
DB	847	CTGGACATCAAGCAGGCGGCCAAGGAGCCCTTCGCGACTACCTGGACCGCTCTCTCAAG	906
QY	901	ACCTCGCGCGGAGCGAGCAGCACCCAGGAGGTGAGAACTGGATGACCGGACACCTGCTG	960
DB			
DB	907	ACCTCGCGCGGAGCGAGCAGCACCCAGGAGGTGAGAACTGGATGACCGGACACCTGCTG	966
QY	961	GTGCAGAAAGCCCAACCCCGACTGCAAGAGCACCATCTCGCGCTCTCGGCGCGGCGGACG	1020
DB			
DB	967	GTGCAGAAAGCCCAACCCCGACTGCAAGAGCACCATCTCGCGCTCTCGGCGCGGCGGACG	1026
QY	1021	CTGGAGGAGATGATGACCGCTGCCAGGCGGTGGGCGGCCAGCCAGCGCGCGGCTG	1080
DB			
DB	1027	CTGGAGGAGATGATGACCGCTGCCAGGCGGTGGGCGGCCAGCCAGCGCGCGGCTG	1086
QY	1081	CTGGCGGAGCGGATGAGCGAGGCCAACCAGCGTGAATGTCAGAGAGCAGCACTTCAGG	1140
DB			
DB	1087	CTGGCGGAGCGGATGAGCGAGGCCAACCAGCGTGAATGTCAGAGAGCAGCACTTCAGG	1146
QY	1141	GGCCCCCGGCGCATCGTCAAGTGCTTCAACTGCGGCAAGGAGGCGCCACATCGCGCGAAC	1200
DB			
DB	1147	GGCCCCCGGCGCATCGTCAAGTGCTTCAACTGCGGCAAGGAGGCGCCACATCGCGCGAAC	1206
QY	1201	TGCGCGCGCCCCCGCAAGAGGGCTGCTGGAAGTGGGCAAGGAGGCGCCACCATGATGAAG	1260
DB			
DB	1207	TGCGCGCGCCCCCGCAAGAGGGCTGCTGGAAGTGGGCAAGGAGGCGCCACCATGATGAAG	1266
QY	1261	GACTGCACGAGCGCGAGGCCAACTTCTTGGGCAAGATCTTGGCCCGCCACAGGCGCGAC	1320
DB			
DB	1267	GACTGCACGAGCGCGAGGCCAACTTCTTGGGCAAGATCTTGGCCCGCCACAGGCGCGAC	1326
QY	1321	CCCGGCAACTTCTTGCAGAGCGCCCCCGAGCCCCCGCCCCCGCGGAGAGCTTCCGCG	1380
DB			
DB	1327	CCCGGCAACTTCTTGCAGAGAGCGCCCCCGAGCCCCCGCCCCCGCGGAGAGCTTCCGCG	1386
QY	1381	TTCGAGGAGACACCCCGCCAGAGAGCAGGAGCAAGGACCGCGGAGACCCCTGACGAGC	1440
DB			
DB	1387	TTCGAGGAGACACCCCGCCAGAGAGCAGGAGCAAGGACCGCGGAGACCCCTGACGAGC	1446
QY	1441	CTGAAGAGCGCTTTCGGCAAGGACCCCTGAGCCAGTAA	1479
DB			
DB	1447	CTGAAGAGCGCTTTCGGCAAGGACCCCTGAGCCAGAA	1485

```

RESULT 3
ACA03592
ID ACA03592 standard; DNA; 4713 BP.
XX
AC ACA03592;
XX
D- 22-MAY-2003 (first entry)
XX
DE Synthetic DNA encoding immunogenic HIV peptide #75.
XX
KW Immunogenic HIV polypeptide; human immunodeficiency virus; HIV;
KW vaccine; gene therapy; packaging cell line; humoral immune response;
KW cellular immune response; gene delivery vector; DNA immunisation;
KW ds.
XX
OS Synthetic.
XX
PN WO2003004657-A1.
XX
PD 16-JAN-2003.
XX
PF 05-JUL-2002; 2002WO-US21421.
XX
PR 05-JUL-2001; 2001US-303192P.
PR 31-AUG-2001; 2001US-316860P.
PR 16-JAN-2002; 2002US-349728P.
PR 16-JAN-2002; 2002US-349793P.
PR 16-JAN-2002; 2002US-349871P.
XX
PA (CHIR ) CHIRON CORP.
XX
PI Zur Megede J, Barnett SW, Lian Y;
XX
DR WPI; 2003-221602/21.
XX
PT New synthetic polynucleotides encoding antigenic HIV type B and/or type
PT C polypeptides, useful as immunogenic compositions or vaccines for
PT generating humoral or cellular immune responses against HIV in a
PT subject, especially humans -
XX
PS Example 1; Fig 80; 262pp; English.
XX
CC The invention describes a synthetic polynucleotide encoding 2 or more
CC immunogenic HIV polypeptides, where at least 2 of the polypeptides are
CC derived from different HIV subtypes. The polynucleotide is useful for
CC immunisation, generation of packaging cell lines, or production of HIV
CC polypeptides. The polynucleotide and its encoded proteins are useful as
CC immunogenic compositions or vaccines for generating humoral or cellular
CC immune responses against HIV in a subject, or for inducing neutralising
CC antibodies against HIV. The gene delivery vector comprising the
CC polynucleotide is also useful for DNA immunisation of, or for
CC generating an immune response (e.g. a humoral or cellular immune
CC response) in, a subject such as a mammal, particularly a human. This
CC sequence encodes a human immunodeficiency virus immunogenic peptide.
XX
SO Sequence 4713 BP; 1041 A; 1682 C; 1462 G; 528 T; 0 other;

Query Match          99.7%; Score 1474.8; DB 25; Length 4713;
Best Local Similarity 99.9%; Pred. NO. 4.7e-186;
Matches 1476; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2   TGGGCGCCGCCAGCATCTGTGGCGGGCGCAAGCTGGACGCCTGGGAGCGCATCGCC 61
DB      1262 TTGGCGCCCGCGCCAGCATCTGTGGCGGGCGCAAGCTGGACGCCTGGGAGCGCATCGCC 1321

QY      62   TGCGCCCCCGCGGCAAGAAGTGTACATGATGAAGCACCTGTGTGGGCCAGCCGCGAGC 121
DB      1322 TGCGCCCCCGCGGCAAGAAGTGCTACATGATGAAGCACCTGTGTGTGGGCCAGCCGCGAGC 1381

QY      122  TGGAGAAGTTGCCCCCTGAACCCCGGCCCTGCTGGAGACCAGCAGGGCTGCAAGCAGATCA 181
DB      1382 TGGAGAAGTTGCCCCCTGAACCCCGGCCCTGCTGGAGACCAGCAGGGCTGCAAGCAGATCA 1441

```


Query Match		99.6%	Score 1473.8;	DB 25;	Length 5184;
Best Local Similarity		99.9%	Pred. No. 6.2e-186;		
Matches 1475;		Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
Qy	3	GGGCGCCCGCGCGCAGCATCCTGGCGCGCGGCAAGCTGGAGCCCTGGAGGSCATCCGCT	62		
Db	1263	GGGCGCCCGCGCGCAGCATCCTGGCGCGCGGCAAGCTGGAGCCCTGGAGGSCATCCGCT	1322		
Qy	63	GGCGCCCGCGCGGCAAGAAGTGCTACATGATGAAGCAACCTGGTGGCGAGCGCGAGCT	122		
Db	1323	GGCGCCCGCGCGGCAAGAAGTGCTACATGATGAAGCAACCTGGTGGCGAGCGCGAGCT	1382		
Qy	123	GGAGAAGTTCCCTGAAACCCCGGCCCTGCTGGAGACCAAGAGGGCTGGAAGCAGATCAT	182		
Db	1383	GGAGAAGTTCCCTGAAACCCCGGCCCTGCTGGAGACCAAGAGGGCTGGAAGCAGATCAT	1442		
Qy	183	CGGCGAGCTGCACCCCGCCCTGCAGACCGGCGAGGAGGAGTGAAGAGCCTGTTCAACAC	242		
Db	1443	CGGCGAGCTGCACCCCGCCCTGCAGACCGGCGAGGAGGAGTGAAGAGCCTGTTCAACAC	1502		
Qy	243	CGTGGCCACCCCTGACTGCGTGCACGAGAGATCGAGGTGCGCGACACCAAGGAGGCCCT	302		
Db	1503	CGTGGCCACCCCTGACTGCGTGCACGAGAGATCGAGGTGCGCGACACCAAGGAGGCCCT	1562		
Qy	303	GGACAAGATCGAGGAGGAGCAGACAAGTGCCAGCAGAGAAGTCCAGCAGCGCGAGGCCGC	362		
Db	1563	GGACAAGATCGAGGAGGAGCAGACAAGTGCCAGCAGAGAAGTCCAGCAGCGCGAGGCCGC	1622		
Qy	363	CGACAAGGGCAAGTGAAGCGAGCAACTACCCCATCGTGCAGAACCTGCAGGCGCAGATGCT	422		
Db	1623	CGACAAGGGCAAGTGAAGCGAGCAACTACCCCATCGTGCAGAACCTGCAGGCGCAGATGCT	1682		
Qy	423	GCACGAGGCCATCAGCCCGCCCGCAGCCCTGAACCCCTGGTGAAGGTGATCCAGGAGAGGC	482		
Db	1683	GCACGAGGCCATCAGCCCGCCCGCAGCCCTGAACCCCTGGTGAAGGTGATCCAGGAGAGGC	1742		
Qy	483	CTTCAGCCCGCAGGTGATCCCATGTTACCCCGCCTGAGCGAGGGCGCCACCCCGCAGGA	542		
Db	1743	CTTCAGCCCGCAGGTGATCCCATGTTACCCCGCCTGAGCGAGGGCGCCACCCCGCAGGA	1802		
Qy	543	CCTGAACACGATGTTGAACACCGTGGCGGCGCACCGGCGCCCATGCGAGTCTGAAGGA	602		
Db	1803	CCTGAACACGATGTTGAACACCGTGGCGGCGCACCGGCGCCCATGCGAGTCTGAAGGA	1862		
Qy	603	CACCATCAACGAGGAGCGCGCGAGTGGAGCGCGTGCACCGCTGCACCGCGCGCAT	662		
Db	1863	CACCATCAACGAGGAGCGCGCGAGTGGAGCGCGTGCACCGCTGCACCGCGCGCAT	1922		
Qy	663	CGCCCGCGCGCGAGATCGCGGAGCGCGCGGCGGCGAGGAGATCGCGGCGCACCGCAT	722		
Db	1923	CGCCCGCGCGCGAGATCGCGGAGCGCGCGGCGGCGAGGAGATCGCGGCGCACCGCAT	1982		
Qy	723	GCAGGAGCAGATCGCCTGGATGACCAAGCAACCCCGCCCAICCCCGTGGCGACATCAAA	782		
Db	1983	GCAGGAGCAGATCGCCTGGATGACCAAGCAACCCCGCCCAICCCCGTGGCGACATCAAA	2042		
Qy	783	GGGTTGGAICATCCTGGGCGCTGACAAAGATCGTGGGATGTACAGCCCGCGTGAAGCATCT	842		
Db	2043	GGGTTGGAICATCCTGGGCGCTGACAAAGATCGTGGGATGTACAGCCCGCGTGAAGCATCT	2102		
Qy	843	GGACATCAAGCAGGCGCCCAAGAGCCCTTCCCGGACTACGTGGACCGCTTCTCAAGAC	902		
Db	2103	GGACATCAAGCAGGCGCCCAAGAGCCCTTCCCGGACTACGTGGACCGCTTCTCAAGAC	2162		
Qy	903	CCTGGCGCGGAGAGAGCACCAGGAGGTGAAGAACTGGATGACCGACACCCCTGCTGCT	962		
Db	2163	CCTGGCGCGGAGAGAGCACCAGGAGGTGAAGAACTGGATGACCGACACCCCTGCTGCT	2222		
Qy	963	GCAGAACGCCAACCCGACTGCAAGACCATCCTGGCGGCTCTGGCGCCCGCGCGAGCCT	1022		
Db	2223	GCAGAACGCCAACCCGACTGCAAGACCATCCTGGCGGCTCTGGCGCCCGCGCGAGCCT	2282		

Qy	1023	GGAGGAGATGATGACCGCCTGCCAGGCGCGTGGGCGCGCCAGCCACAGGSCCGCGTCT	1082		
Db	2283	GGAGGAGATGATGACCGCCTGCCAGGCGCGTGGGCGCGCCAGCCACAGGSCCGCGTCT	2342		
Qy	1083	GGCCGAGCGGATGAGCCAGSCCAACACCAAGCGTGAIGATGCAGAGAGCAACTTCAAGGG	1142		
Db	2343	GGCCGAGCGGATGAGCCAGSCCAACACCAAGCGTGAIGATGCAGAGAGCAACTTCAAGGG	2402		
Qy	1143	CCCCGGCGCATCGTCAAGTGTCTCAACTGCGGCAAGGAGGGCCACATGCGCCGCACTG	1202		
Db	2403	CCCCGGCGCATCGTCAAGTGTCTCAACTGCGGCAAGGAGGGCCACATGCGCCGCACTG	2462		
Qy	1203	CCGCGCCCGCCGCAAGAAGGCGTCTGGAAGTGGGCAAGGAGGGCCACAGATGAAGGA	1262		
Db	2463	CCGCGCCCGCCGCAAGAAGGCGTCTGGAAGTGGGCAAGGAGGGCCACAGATGAAGGA	2522		
Qy	1263	CTGCACCGAGCGCCAGGCCAACTTCTGGCAAGATCTGGCCAGCCAGCCAGGGCCGCC	1322		
Db	2523	CTGCACCGAGCGCCAGGCCAACTTCTGGCAAGATCTGGCCAGCCAGCCAGGGCCGCC	2582		
Qy	1323	CGCAACTTCTGCAGAGCGCCCGGAGCCACCGCCCGCCCGCGGAGAGCTTCCGCTT	1382		
Db	2583	CGCAACTTCTGCAGAGCGCCCGGAGCCACCGCCCGCCCGCGGAGAGCTTCCGCTT	2642		
Qy	1383	CGAGGAGACCCCGCCCGGCGAGAGCAAGGAGCAAGGAGCCGCGAGACCTGACCACT	1442		
Db	2643	CGAGGAGACCCCGCCCGGCGAGAGCAAGGAGCAAGGAGCCGCGAGACCTGACCACT	2702		
Qy	1443	GAAGAGCCTGTGGCAACGACCCCGCTGAGCCAGTAA	1479		
Db	2703	GAAGAGCCTGTGGCAACGACCCCGCTGAGCCAGTAA	2739		
RESULT 10					
ACA03590					
ID	ACA03590 standard; DNA; 2742 BP.				
XX	ACA03590;				
AC					
DT	22-MAY-2003 (first entry)				
XX	Synthetic DNA encoding immunogenic HIV peptide #73.				
DE	immunogenic HIV polypeptide; human immunodeficiency virus; HIV;				
KW	vaccine; gene therapy; packaging cell line; humoral immune response;				
KW	cellular immune response; gene delivery vector; DNA immunisation;				
KW	ds.				
XX	Synthetic.				
GS	WO2003004657-A1.				
XX	16-JAN-2003.				
PN	05-JUL-2002; 2002WO-US21421.				
XX	05-JUL-2001; 2001US-303192P.				
PR	31-AUG-2001; 2001US-316860P.				
PR	16-JAN-2002; 2002US-349728P.				
PR	16-JAN-2002; 2002US-349793P.				
PR	16-JAN-2002; 2002US-349871P.				
XX	(CHIR) CHIRON CORP.				
PA	zur Megede J, Barnett SW, Lian Y;				
XX	WPI; 2003-221602/21.				
DR	New synthetic polynucleotides encoding antigenic HIV type B and/or type				
XX	C polypeptides, useful as immunogenic compositions or vaccines for				
PT	generating humoral or cellular immune responses against HIV in a				
PT	subject, especially humans -				
XX					

PS Example 1; Fig 78; 262pp; English.

XX The invention describes a synthetic polynucleotide encoding 2 or more
CC immunogenic HIV polypeptides, where at least 2 of the polypeptides are
CC derived from different HIV subtypes. The polynucleotide is useful for
CC immunisation, generation of packaging cell lines, or production of HIV
CC polypeptides. The polynucleotide and its encoded proteins are useful as
CC immunogenic compositions or vaccines for generating humoral or cellular
CC immune responses against HIV in a subject, or for inducing neutralising
CC antibodies against HIV. The gene delivery vector comprising the
CC polynucleotide is also useful for DNA immunisation of, or for
CC generating an immune response (e.g. a humoral or cellular immune
CC response) in, a subject such as a mammal, particularly a human. This
CC sequence encodes a human immunodeficiency virus immunogenic peptide.
XX
SQ Sequence 2742 BP; 578 A; 1020 C; 859 G; 285 T; 0 other;

Query Match 95.6%; Score 1472.4; DB 25; Length 2742;
Best local Similarity 99.9%; Pred. No. 1e-185;
Matches 1473; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TGGGCGCCGCGCAGCATCTGCGCGGCGGCAAGCTGACGCTGGGAGCGCATCCGCC 41
DB 1262 TCGGCGCCGCGCAGCATCTGCGCGGCGGCAAGCTGACGCTGGGAGCGCATCCGCC 1321
QY 62 TGGGCGCCGCGGCGCAGCATCTGCGCGGCGGCAAGCTGACGCTGGGAGCGCATCCGCC 121
DB 1322 TCGGCGCCGCGGCGCAGCATCTGCGCGGCGGCAAGCTGACGCTGGGAGCGCATCCGCC 1381
QY 122 TGGGCGCGGCGGCGCAGCATCTGCGCGGCGGCAAGCTGACGCTGGGAGCGCATCCGCC 181
DB 1382 TGGGCGCGGCGGCGCAGCATCTGCGCGGCGGCAAGCTGACGCTGGGAGCGCATCCGCC 1441
QY 182 TCGGCGCGGCGGCGCAGCATCTGCGCGGCGGCAAGCTGACGCTGGGAGCGCATCCGCC 241
DB 1442 TCGGCGCGGCGGCGCAGCATCTGCGCGGCGGCAAGCTGACGCTGGGAGCGCATCCGCC 1501
QY 242 CCGTGGCGCCGCGGCGCAGCATCTGCGCGGCGGCAAGCTGACGCTGGGAGCGCATCCGCC 301
DB 1502 CCGTGGCGCCGCGGCGCAGCATCTGCGCGGCGGCAAGCTGACGCTGGGAGCGCATCCGCC 1561
QY 302 TGGGCGCGGCGGCGCAGCATCTGCGCGGCGGCAAGCTGACGCTGGGAGCGCATCCGCC 361
DB 1562 TGGGCGCGGCGGCGCAGCATCTGCGCGGCGGCAAGCTGACGCTGGGAGCGCATCCGCC 1621
QY 362 CCGGCGCGGCGGCGCAGCATCTGCGCGGCGGCAAGCTGACGCTGGGAGCGCATCCGCC 421
DB 1622 CCGGCGCGGCGGCGCAGCATCTGCGCGGCGGCAAGCTGACGCTGGGAGCGCATCCGCC 1681
QY 422 TGGGCGCGGCGGCGCAGCATCTGCGCGGCGGCAAGCTGACGCTGGGAGCGCATCCGCC 481
DB 1682 TGGGCGCGGCGGCGCAGCATCTGCGCGGCGGCAAGCTGACGCTGGGAGCGCATCCGCC 1741
QY 482 CCGTGGCGCCGCGGCGCAGCATCTGCGCGGCGGCAAGCTGACGCTGGGAGCGCATCCGCC 541
DB 1742 CCGTGGCGCCGCGGCGCAGCATCTGCGCGGCGGCAAGCTGACGCTGGGAGCGCATCCGCC 1801
QY 542 ACCTGAACACGATGTTGAACACCGCTGGGCGGCGGCAAGCTGACGCTGGGAGCGCATCCGCC 601
DB 1802 ACCTGAACACGATGTTGAACACCGCTGGGCGGCGGCAAGCTGACGCTGGGAGCGCATCCGCC 1861
QY 602 ACACCATCAACGAGGAGGCGCGGAGTGGACCGCGCTGCACCCCGTGCACGCGCGGCCCA 661
DB 1862 ACACCATCAACGAGGAGGCGCGGAGTGGACCGCGCTGCACCCCGTGCACGCGCGGCCCA 1921
QY 662 TCGGCGCCGCGGCGGCGCAGCATCTGCGCGGCGGCAAGCTGACGCTGGGAGCGCATCCGCC 721
DB 1922 TCGGCGCCGCGGCGGCGCAGCATCTGCGCGGCGGCAAGCTGACGCTGGGAGCGCATCCGCC 1981
QY 722 TGCAGGAGCGAGATCGCCCTGGATGACGACGACGACGACGACGACGACGACGACGACGAC 781
DB 1982 TGCAGGAGCGAGATCGCCCTGGATGACGACGACGACGACGACGACGACGACGACGACGAC 2041

QY 782 AGCGGTGGATCATCTCTGGGCTGAACAAGATCTGCGGATGTACAGCCCGGTGAGCATCC 841
DB 2042 AGCGGTGGATCATCTCTGGGCTGAACAAGATCTGCGGATGTACAGCCCGGTGAGCATCC 2102
QY 842 TGGACATCAAGCAGGCGGCGGCAAGGAGGCGCTTCCGCGACTACGTTGGACCGCTTCTCAAGA 901
DB 2102 TGGACATCAAGCAGGCGGCGGCAAGGAGGCGCTTCCGCGACTACGTTGGACCGCTTCTCAAGA 2161
QY 902 CCGTGGCGCCGCGGCGCAGCAGGACCCAGGAGGTGAAGAATCTGATGACGACACCCCTGCTGG 961
DB 2162 CCGTGGCGCCGCGGCGCAGCAGGACCCAGGAGGTGAAGAATCTGATGACGACACCCCTGCTGG 2221
QY 962 TGCAGAAAGCGCAACCCCGACTGCAAGACCATCTGCGGCTCTCGGCGCGGCGGCGGCGGCGG 1021
DB 2222 TGCAGAAAGCGCAACCCCGACTGCAAGACCATCTCTGCGGCTCTCGGCGCGGCGGCGGCGGCGG 2281
QY 1022 TGGAGGAGATGATGACCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1081
DB 2282 TGGAGGAGATGATGACCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2341
QY 1082 TGGCGGAGCGGATGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1141
DB 2342 TGGCGGAGCGGATGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2401
QY 1142 GCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1201
DB 2402 GCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2461
QY 1202 GCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1261
DB 2462 GCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2521
QY 1262 ACTGACCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1321
DB 2522 ACTGACCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2581
QY 1322 CCGGCG 1381
DB 2582 CCGGCG 2641
QY 1382 TCGAGGAGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1441
DB 2642 TCGAGGAGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2701
QY 1442 TGAAGAGCGCTGTTGCGGCAACGACCCCTGAGGCCA 1475
DB 2702 TGAAGAGCGCTGTTGCGGCAACGACCCCTGAGGCCA 2735

RESULT 11

AAA51625

15 AAA51625 standard; DNA; 1479 BP.

XX AAA51625;

XX 31-OCT-2000 (first entry)

XX HIV codon-optimized synthetic Gag polynucleotide.

XX Gag; expression cassette; antigenic; type C; HIV; Env; synthetic;
KW DNA immunization; packaging cell line; antigen presentation; ss.

XX Human immunodeficiency virus type C strain AF110965.
OS Synthetic.

XX WO200039304-A2.

XX 06-JUL-2000.

XX 30-DEC-1999; 99WO-US31273.

XX 31-DEC-1998; 98US-0114495.

PR 01-SEP-1999; 99US-0152195.

XX (CHIR) CHIRON CORP.
PA Barnett S, Zur Megede J;
XX WPI; 2000-452401/39.
DR
XX Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV
PT Env polypeptide and the polypeptide useful for immunizing a mammal
PT especially human against HIV
XX
PS Disclosure: Page 103-104; 113pp; English.
XX
XX Expression cassettes comprising a polynucleotide encoding antigenic
CC type C human immunodeficiency virus (HIV) Gag or Env polypeptides are
CC useful in DNA immunization, generation of packaging cell lines and
CC production of Gag- and/or Env-containing proteins. Synthetic Env and Gag
CC expression cassettes exhibit increased potency for induction of
CC cytotoxic T-lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1
CC self-assemble into non-infectious virus-like particles which are used as
CC a matrix for the proper presentation of an antigen entrapped or
CC associated to the immune system of the host.
XX
SQ Sequence 1479 BP; 325 A; 533 C; 461 G; 160 T; 0 other;

Query Match 98.9%; Score 1463; DB 21; Length 1479;
Best local Similarity 99.3%; Pred. No. 2e-184;
Matches 1469; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 ATGGGCGCCCGCGGCGAGCATCTCTGGCGGCGGGAAGCTGGAGCGCTGGGAGCGATCGGC 60
DB 1 ATGGGCGCCCGCGGCGAGCATCTCTGGCGGCGGGAAGCTGGAGCGCTGGGAGCGATCGGC 60

QY 61 CTGGCGCGCGCGCGGCGAAGAAGTGTCTACATGATGAAGCACCTGGTGTGGCGAGCGCGGAG 120
DB 61 CTGGCGCGCGCGCGGCGAAGAAGTGTCTACATGATGAAGCACCTGGTGTGGCGAGCGCGGAG 120

QY 121 CTGGAGAAGTTCGCCTGAACCCCGGCTGTGGAGACCAAGCTGGAGGCTGCAAGCAGATC 180
DB 121 CTGGAGAAGTTCGCCTGAACCCCGGCTGTGGAGACCAAGCTGGAGGCTGCAAGCAGATC 180

QY 181 ATCCGCGAGCTGCACCCCGCCCTGCAGACCGCGGAGGAGGAGCTGAAGAGCTGTTCAC 240
DB 181 ATCCGCGAGCTGCACCCCGCCCTGCAGACCGCGGAGGAGGAGCTGAAGAGCTGTTCAC 240

QY 241 ACCGTGGCCACCTGTACTGCTGACGAGAGATCGAGGTCCGGACACCAAGAGGCG 300
DB 241 ACCGTGGCCACCTGTACTGCTGACGAGAGATCGAGGTCCGGACACCAAGAGGCG 300

QY 301 CTGGACAAGATCGAGGAGGAGGAGCAACAAGTGCACGAGAGATCCAGAGCGCGAGGCG 360
DB 301 CTGGACAAGATCGAGGAGGAGGAGCAACAAGTGCACGAGAGATCCAGAGCGCGAGGCG 360

QY 361 GCGGACAAGGCAAGGTGAGCCCAACTACCCCATCTGTCAGAACCTGACGCGGCGAGATG 420
DB 361 GCGGACAAGGCAAGGTGAGCCCAACTACCCCATCTGTCAGAACCTGACGCGGCGAGATG 420

QY 421 GTGCACGAGGCCATCAGCCCGCCGACCCCTGAACGCTGGTGAAGGTGATCGAGGAGAG 480
DB 421 GTGCACGAGGCCATCAGCCCGCCGACCCCTGAACGCTGGTGAAGGTGATCGAGGAGAG 480

QY 481 GCCTTCAGCCCGGAGGTGATCCCATGTTTACCGCCCTGAGCGAGGCGGCGACCCCGCAG 540
DB 481 GCCTTCAGCCCGGAGGTGATCCCATGTTTACCGCCCTGAGCGAGGCGGCGACCCCGCAG 540

QY 541 GACCTGAACACGATGTTGAACACCGTGGCGGCGCACCGAGCGGCCCATGTCAGATGCTGAAG 600
DB 541 GACCTGAACACGATGTTGAACACCGTGGCGGCGCACCGAGCGGCCCATGTCAGATGCTGAAG 600

QY 601 GACACCATCAACGAGGAGGCGCGGAGTGGGAGCGGTGCACCCCGTGCACCGCGCGCC 660
DB 601 GACACCATCAACGAGGAGGCGCGGAGTGGGAGCGGTGCACCCCGTGCACCGCGCGCC 660

QY 661 ATGCGCCCGCGGCGAGATGCGCGGAGCGCGCGGCGAGCGACATCGCGGCGACCGACGACC 720
DB 661 ATGCGCCCGCGGCGAGATGCGCGGAGCGCGCGGCGAGCGACATCGCGGCGACCGACGACC 720

QY 721 CTGCAGGAGCAGATCGCTGGATGACCAACCAACCCCGCATCCCGTGGCGGACATCTAC 780
DB 721 CTGCAGGAGCAGATCGCTGGATGACCAACCAACCCCGCATCCCGTGGCGGACATCTAC 780

QY 781 AAGCGGTGGATCATCTCTGGGCTGAACAGAGATGCTGGGAGTACACCCCGGTGAGCATC 840
DB 781 AAGCGGTGGATCATCTCTGGGCTGAACAGAGATGCTGGGAGTACACCCCGGTGAGCATC 840

QY 841 CTGGACATCAAGCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
DB 841 CTGGACATCAAGCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900

QY 901 ACCCTGGCGGCGGAGCAGACACCCAGGAGGTGAAGAGTGGATGACCGGACACCTGCTG 960
DB 901 ACCCTGGCGGCGGAGCAGACACCCAGGAGGTGAAGAGTGGATGACCGGACACCTGCTG 960

QY 961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCTCTGGCGGCTCTCGGCGGCGGCGGCGG 1020
DB 961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCTCTGGCGGCTCTCGGCGGCGGCGGCGG 1020

QY 1021 CTGGAGGAGATGATGACCCGCTGCCAGGCGGTGGGCGGCGGCGGCGGCGGCGGCGGCG 1080
DB 1021 CTGGAGGAGATGATGACCCGCTGCCAGGCGGTGGGCGGCGGCGGCGGCGGCGGCGGCG 1080

QY 1081 CTGGCGGAGGCGGATGAGCCAGGCGGCAACACCGGCTGATGCGAGAGAGCAACTTCAAG 1140
DB 1081 CTGGCGGAGGCGGATGAGCCAGGCGGCAACACCGGCTGATGCGAGAGAGCAACTTCAAG 1140

QY 1141 GCGCGCGGCGGCGGATGCTCAAGTGTCTCAACTGCGGCAAGGAGGCGGCGGCGGCGGCGG 1200
DB 1141 GCGCGCGGCGGCGGATGCTCAAGTGTCTCAACTGCGGCAAGGAGGCGGCGGCGGCGGCGG 1200

QY 1201 TGCGCGCGCGCGCGGCAAGAGGCTGCTGGAGTGGGCAAGGAGGCGGCGGCGGCGGCGG 1260
DB 1201 TGCGCGCGCGCGCGGCAAGAGGCTGCTGGAGTGGGCAAGGAGGCGGCGGCGGCGGCGG 1260

QY 1261 GACTGCACCGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1320
DB 1261 GACTGCACCGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1320

QY 1321 CCGGCGCAACTCTCTGCAGAGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1380
DB 1321 CCGGCGCAACTCTCTGCAGAGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1380

QY 1381 TCGAGGAGACCAACCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1440
DB 1381 TCGAGGAGACCAACCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1440

QY 1441 CTGAAGAGCGCTCTCTCGGCAACGACCCCGCTGAGCGGAGTAA 1479
DB 1441 CTGAAGAGCGCTCTCTCGGCAACGACCCCGCTGAGCGGAGTAA 1479

RESULT 12
ABL39957
ID ABL39957 standard; DNA: 1479 BP.
XX
AC ABL39957;
XX
DT 15-MAY-2002 (first entry)
XX
DE Synthetic Gag polynucleotide sequence SEQ ID NO:20.
XX
KW Human immunodeficiency virus type C; antigenic HIV type C protein;
KW immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
KW immunostimulant; gene therapy; gene; ds.
OS Human immunodeficiency virus type C.
OS Synthetic.

XX 08-NOV-2002 (first entry)
XX HIV-1 p55gag polypeptide coding sequence 3.
XX
KW HIV; ds: vaccine; gene; immune response; microparticle;
KW adsorbent surface; poly(alpha-hydroxy acid); polyhydroxy butyric acid;
KW polycaprolactone; polyorthoester; polycyanoacrylate; detergent;
KW submicron emulsion; viral infection; bacterial infection;
KW parasitic infection; HIV-1 p55gag polypeptide.
XX
OS Human immunodeficiency virus type 1.
XX
FH Key Location/Qualifiers
FT variation replace (282, C)
FT /*tag= a
FT variation replace (552, G)
FT /*tag= b
FT variation replace (555, T)
FT /*tag= c
FT variation replace (786, G)
FT /*tag= d
FT variation replace (813, G)
FT /*tag= e
FT variation replace (1002, T)
FT /*tag= f
FT variation replace (1005, C)
FT /*tag= g
FT variation replace (1092, G)
FT /*tag= h
FT variation replace (1149, G)
FT /*tag= i
FT variation replace (1158, C)
FT /*tag= j
XX
PN WC2002262C9-A2.
XX
PD 04-APR-2002.
XX
PF 28-SEP-2001; 2031WQ-US30540.
XX
PR 28-SEP-2000; 2000US-236105P.
PR 30-AUG-2001; 2001US-315905P.
XX
PA (CHIR) CHIRON CORP.
XX
PI O'hagan D, Otten G, Donnelly CJ, Polo JM, Barnett S, Singh M;
PI Ulmer J, Dubensky TW;
XX
XX WPI; 2002-519084/55.
XX
XX A microparticle to which a biologically active macromolecule is
PT adsorbed, for use as a vaccine composition to treat viral, bacterial or
PT parasitic infections, comprises a polymer microparticle, a detergent
PT and a submicron emulsion -
XX
PS Claim 72; Fig 5; 100pp; English.
XX
CC The invention relates to a method of raising an immune response in a host
CC animal. The method of the invention comprises administering a
CC microparticle that has an adsorbent surface to which a first biologically
CC active macromolecule (e.g. a nucleic acid) has been adsorbed. The
CC microparticle comprises a polymer microparticle of poly(alpha-hydroxy
CC acid), a polyhydroxy butyric acid, a polycaprolactone, a polyorthoester,
CC a polycyanoacrylate, a detergent, and submicron emulsion. The method/
CC microparticle of the invention is useful for immunising a host animal
CC against viral, bacterial or parasitic infections. The present DNA
CC sequence encodes a HIV-1 p55gag polypeptide.
XX
SQ Sequence 1479 BP: 325 A; 534 C; 460 G; 160 T; 0 other;

Query Match 98.8%; Score 1461.4; DB 24; Length 1479;
Best Local Similarity 99.3%; Pred. No. 3.2e-184;

Matches 1468: Conservative 0; Mismatches 11; Indels 0; Gaps 0;				
QY	1	ATGGGCGCGCGCGCAGCATCTTGGCGCGCGGCAAGCTGGAGGCTGGGAGGCGCATCGCG	60	
DB				
QY	61	CTGGCGCGCGCGCGCAAGAGTGTATCATGATGAAGCACCTTGTGTGGCGAGCGCGGAG	120	
DB				
QY	121	CTGGAGAAAGTTCGCCCTGAACCCCGCCCTGCTGGAGACCAAGCGAGGGTGAAGAGCATC	180	
DB				
QY	181	ATCCGCCAGCTGCACCCCGCCCTGCAGACCGGAGCGGAGAGCTGAAGAGGCTGTTCAC	240	
DB				
QY	241	ACCGTGGCCACCCCTGTACTGCTGCACGAGAGATCGAGGTCCGGACACCAAGGAGGCC	300	
DB				
QY	301	CTGGACAAGATCGAGGAGGAGCAGAACAAAGTCCAGCAGAGATCCAGAGCGCGAGGCC	360	
DB				
QY	361	GCCGACAAGGGCAAGGTGAGCCAGAACTACCCATCGTGCAGAACCTGCAGGCGCAGATG	420	
DB				
QY	421	GTGCACAGGGCCATCAGCCCCCGCCAGCTGAACGCCCTGCTGAAGGTGATCGAGGAGAAG	480	
DB				
QY	481	GCCTTCAGCCCCGAGGTGATCCCATGTTTACCGCCCTGAGCGAGGGCGGCGACCCCCCAG	540	
DB				
QY	541	GACCTGAACACGATGTTGAACACCGTGGGCGGCGCACCCAGGCCCATGCAGATGCTGAAG	600	
DB				
QY	601	GACACCATCAACGAGGAGCGCGCGAGTGGGACCGGTGCACCGCGTGCACCGCGGCC	660	
DB				
QY	661	ATCGCCCCCGCCAGATCGCGAGCGCGCGCGGAGATCGCGGCGGAGATCGCGGCGGCGC	720	
DB				
QY	721	CTGCAGGAGCAGATCGCTGGATGACCAAGAGTGTGGGATGTACAGCCCGGTGAGCATC	780	
DB				
QY	781	AAGCGGTGGATCATCTCTGGCGCTGAACAAAGTGTGGGATGTACAGCCCGGTGAGCATC	840	
DB				
QY	841	CTGGACATCAAGCAGGCGCCCAAGAGCGCCCTTCGCGGACTACGTGGACCGCTTCTTCAAG	900	
DB				
QY	901	ACCGTGGCGCGCGGAGCAGACCGAGGAGTGAAGAACTGGATGACCGACACCGCTGCTG	960	
DB				
QY	961	GTGCAGAACGCCAACCCCGACTGCAAGACCATCTCTGCGGCGTCTCGGCGCGCGCGCAGC	1020	
DB				
QY	1021	CTGGAGGAGATGATGACCGCCCTGCCAGGGCGTGGGCGGCGCGGCGGCGGCGCGGCGTG	1080	
DB				
QY	1021	CTGGAGGAGATGATGACCGCCCTGCCAGGGCGTGGGCGGCGCGGCGGCGGCGGCGGCGTG	1080	
DB				

Db 718 CTGCAGGAGAGATCGCCTGGATGACCAAGCAACCCCCCATCCCCGCGGAGGACATCTAC 777
QY 781 AAGCGGTGGATCAICCTGGCCCTGAACRAGATCGTGGGATGTACAGCCCCCGGAGCATC 840
Db 778 AAGCGCTGGATCATCCTGGCCCTGAACRAGATCGTGGCATGTACAGCCCCCGTGAAGCATC 837
QY 841 CTGGACATCAAGCAGGCGGCCCAAGGAGGCCCTCCGCGACTACGTGGACCGCTTCCTCAAG 900
Db 838 CTGGACATCAAGCAGGCGGCCCAAGGAGGCCCTCCGCGACTACGTGGACCGCTTCCTCAAG 897
QY 901 ACCCTGCGGCGCGAGCAGACACCCAGGAGGTGAAGAACTGGATGACGGACACCCCTGCTG 960
Db 898 ACCCTGCGGCGCGAGCAGACACCCAGGAGGTGAAGAACTGGATGACGGACACCCCTGCTG 957
QY 961 GTGCAGAACGGCAACCCGACTGCAAGACATCCTGGCGCTCTCGGCCCGGGGCGCAGC 1020
Db 958 GTGCAGAACGGCAACCCGACTGCAAGACATCCTGGCGCTCTCGGCCCGGGGCGCAGC 1017
QY 1021 CTGGAGGAGATGATACCGCTGCCAGGGGCTGGCGGCCCGCAGCCCAAGGCCCGCTG 1080
Db 1018 CTGGAGGAGATGATACCGCTGCCAGGGGCTGGCGGCCCGCAGCCCAAGGCCCGCTG 1077
QY 1081 CTGGCGGAGGGATGAGCCAGGGCCAAACACACGCGTGTATGTCAGAAAGAGCAGACCTCAAG 1140
Db 1078 CTGGCGGAGGGCATGAGCCAGGGCCAAACACGCAACATCCTGGTGCAGCGGAGCAACTCAAG 1137
QY 1141 GGGCCCCCGGCGCATCGTCAAGTGTCTCAACTCGGGCAAGGGGCGCACATGGCCCCGAAG 1200
Db 1138 GGCAGCAACCGCATCATCAAGTGTCTCAACTCGGGCAAGGTGGGCCACATGGCCCCGAAC 1197
QY 1201 TCCCGCGCCCGCCGCAAGAGGGGCTGCTGAAGTGGGCAAGGAGGGGCGCACAGATGAAG 1260
Db 1198 TGGCGCGCCCGCCGCAAGAGGGGCTGCTGAAGTGGGCAAGGAGGGGCGCACAGATGAAG 1257
QY 1261 GACTGCACCGAGCGCCGAGGCGCAACTTCCTGGGCAAGATCTGGCCCGAGCCACAGGCGCGC 1320
Db 1258 GACTGCACCGAGCGCCGAGGCGCAACTTCCTGGGCAAGATCTGGCCCGAGCCACAGGCGCGC 1317
QY 1321 CCGGSGCAACTTCCTGCAGAGCGCGC-----CCGAGCGCCAGCGCT 1359
Db 1318 CCGGSGCAACTTCCTGCAGAGCGCCCGGAGCGCCCGGAGCGCCAGCGCGC 1377
QY 1360 CCGCCCGCGCGAGAGCTTCGCGCTCGAGGAGACACCGCCCGCCAGAGAGGAGGCAAG 1419
Db 1378 CCGCCCGCGCGAGAGCTTCGCGCTCGAGGAGACACCGCCCGCTGCGCCGCAAGGAG---AAG 1434
QY 1420 GAGCGGAGAGCGCTGACCGAGCGCTGAAGAGCGTGTTCGGCAAGACCGCCCTGAGCCAGTAA 1479
Db 1435 GAGCGGAGCGCGCTGACCGAGCGCTGAAGAGCGTGTTCGGCAAGACCGCCCTGAGCCAGTAA 1494

CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)

FEATURES		Location/Qualifiers	
source	1..1479	/organism="synthetic construct"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:32630"	
		/note="synthetic Gag of HIV strain AF110965"	
BASE COUNT		325 a	529 c 463 g 162 t
ORIGIN			
Query Match 100.0%; Score 1479; DB 6; Length 1479;			
Best Local Similarity 100.0%; Pred. No. 3e-151;			
Matches 1479; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	ATGGGGCGCGCCGACCATCTCTGGGGGGGCAAGCTGGACGCCCTGGGAGCGCATCCGC	60
Db	1	ATGGGGCGCGCGCCGACCATCTCTGGGGGGGCAAGCTGGACGCCCTGGGAGCGCATCCGC	60
QY	61	CTGGCGCCCGCGGCGCAAGAAGTGCTACATGATGAAGACCTGGTGTGGGCCAGCCCGGAG	120
Db	61	CTGCGCCCGCGGCGCAAGAAGTGCTACATGATGAAGACCTGGTGTGGGCCAGCCCGGAG	120
QY	121	CTGGAGAGTTGCGCCCTGAACCCCGGCTGCTGGAGACCGAGGCGTGAAGCAGATC	180
Db	121	CTGGAGAGTTGCGCCCTGAACCCCGGCTGCTGGAGACCGAGGCGTGAAGCAGATC	180
QY	181	ATCCGCCAGCTGCACCCCGCCCTGCAGACCGGCGAGCGAGAGCTGAAGAGCCCTGTCAC	240
Db	181	ATCCGCCAGCTGCACCCCGCCCTGCAGACCGGCGAGCGAGAGCTGAAGAGCCCTGTCAC	240
QY	241	ACCGTGGCCACCCCTGTACTGCGTGCCACGAGAAGATCGAGGTCCGGACACCAAGGAGCC	300
Db	241	ACCGTGGCCACCCCTGTACTGCGTGCCACGAGAAGATCGAGGTCCGGACACCAAGGAGCC	300
QY	301	CTGGACAAGATCGAGGAGGAGCAGAACAGTGGCAGCAGAGATCCAGCAGGCGCGAGGCC	360
Db	301	CTGGACAAGATCGAGGAGGAGCAGAACAGTGGCAGCAGAGATCCAGCAGGCGCGAGGCC	360
QY	361	GCCGACAAGGGCAAGGTGAGCCAGAACTACCCCATCTGCGAGAACCCTGCAGGSCCAGATG	420
Db	361	GCCGACAAGGGCAAGGTGAGCCAGAACTACCCCATCTGCGAGAACCCTGCAGGSCCAGATG	420
QY	421	GTGCACCAAGGCCATCAGCCCCCGCACCCCTGAACGCCCTGGGTAAGGTGATCGAGGAGAG	480
Db	421	GTGCACCAAGGCCATCAGCCCCCGCACCCCTGAACGCCCTGGGTAAGGTGATCGAGGAGAG	480
QY	481	GCCTTCAGCCCGGAGGTGATCCCATGTTACCGGCCCTGAGCGAGGGGCCACCCCGGAG	540
Db	481	GCCTTCAGCCCGGAGGTGATCCCATGTTACCGGCCCTGAGCGAGGGGCCACCCCGGAG	540
QY	541	GACCTGAACAGCATGTTGAACACCGTGGGCGGCCACCGAGGCGCCCATSCAGATGCTGAG	600
Db	541	GACCTGAACAGCATGTTGAACACCGTGGGCGGCCACCGAGGCGCCCATSCAGATGCTGAG	600
QY	601	GACACCATCAACGAGGAGGCGCGGAGTGGGACCGCGTGCACCCCGTGCACGCGCGGCGCC	660
Db	601	GACACCATCAACGAGGAGGCGCGGAGTGGGACCGCGTGCACCCCGTGCACGCGCGGCGCC	660
QY	661	ATCGCCCCCGCCAGATGCGCGAGGCCCGCGGACGATCGCCGSCACCAACCGAGACC	720
Db	661	ATCGCCCCCGCCAGATGCGCGAGGCCCGCGGACGATCGCCGSCACCAACCGAGACC	720
QY	721	CTGCAGGAGCAGATCGCCTGGATGACCAAGCAACCCCGCATCCCGTGGGCGACATCTAC	780
Db	721	CTGCAGGAGCAGATCGCCTGGATGACCAAGCAACCCCGCATCCCGTGGGCGACATCTAC	780
QY	781	AAGCGGTGGATCATCTCTGGGCTGAACAGATCGTGGGATGTACAGCCCGCTGAGCATC	840
Db	781	AAGCGGTGGATCATCTCTGGGCTGAACAGATCGTGGGATGTACAGCCCGCTGAGCATC	840
QY	841	CTGGACATCAACGAGGCGCCCAAGGAGCCCTTCCGGCGACTACGTGGAGCGCTTCTTCAAG	900
Db	841	CTGGACATCAACGAGGCGCCCAAGGAGCCCTTCCGGCGACTACGTGGAGCGCTTCTTCAAG	900

QY	901	ACCTTGGCGCGCGGAGCAGACACCCAGGAGTGAAGAACTGGATGACCGACACCCCTGCTG	960
Db	901	ACCTTGGCGCGCGGAGCAGACACCCAGGAGTGAAGAACTGGATGACCGACACCCCTGCTG	960
QY	961	GTGCAGAAAGCGCAACCCCGACTGCAAGACCATCTGCGCGCTCTCGGCCCGGCGCCAGC	1020
Db	961	GTGCAGAAAGCGCAACCCCGACTGCAAGACCATCTGCGCGCTCTCGGCCCGGCGCCAGC	1020
QY	1021	CTGGAGGAGATGATGACCGCTGCCAGGGCGTGGCGGCCAGCCACAAAGGCCCGCGTG	1080
Db	1021	CTGGAGGAGATGATGACCGCTGCCAGGGCGTGGCGGCCAGCCACAAAGGCCCGCGTG	1080
QY	1081	CTGGCCGAGGGGATGAGCCAGGCCAACACACAGCGTGATGATGCAGAAGAGCAACTTCAAG	1140
Db	1081	CTGGCCGAGGGGATGAGCCAGGCCAACACACAGCGTGATGATGCAGAAGAGCAACTTCAAG	1140
QY	1141	GGCCCCCGCGCATGCTCAAGTGTCTCAACTGCGGCAAGGAGGGCCACATGCGCCGCAAC	1200
Db	1141	GGCCCCCGCGCATGCTCAAGTGTCTCAACTGCGGCAAGGAGGGCCACATGCGCCGCAAC	1200
QY	1201	TGCGCGCCCGCCGCAAGAGGGCTGCTGGAAGTGGCGGCAAGGAGGGCCACAGATGAAG	1260
Db	1201	TGCGCGCCCGCCGCAAGAGGGCTGCTGGAAGTGGCGGCAAGGAGGGCCACAGATGAAG	1260
QY	1261	GACTGCACCGAGCGCCAGGCCAACTTCTTGGCAAGATCTGGCCCCAGCCCAAGGGCCGC	1320
Db	1261	GACTGCACCGAGCGCCAGGCCAACTTCTTGGCAAGATCTGGCCCCAGCCCAAGGGCCGC	1320
QY	1321	CGGGCAACTTCTTGCAGAGCGCCCGGAGCCACCGCCCCCGCCCGCGGAGAGCTTCCGC	1380
Db	1321	CGGGCAACTTCTTGCAGAGCGCCCGGAGCCACCGCCCCCGCCCGCGGAGAGCTTCCGC	1380
QY	1381	TTCCAGGAGACACACCCCGCGGCAAGAGCAGGAGCAAGGACCGCGAGACCTTGACCAGC	1440
Db	1381	TTCCAGGAGACACACCCCGCGGCAAGAGCAGGAGCAAGGACCGCGAGACCTTGACCAGC	1440
QY	1441	CTGAAGAGCCCTTTCGGCAACGACCCCTTGAGCCAGTAA	1479
Db	1441	CTGAAGAGCCCTTTCGGCAACGACCCCTTGAGCCAGTAA	1479
RESULT 2			
AX468543			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
BASE COUNT			
ORIGIN			
Query Match 100.0%; Score 1479; DB 6; Length 1479;			
Best Local Similarity 100.0%; Pred. No. 3e-151;			
Matches 1479; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	ATGGGGCGCGCGCCGACCATCTCTGGGGGGGCAAGCTGGACGCCCTGGGAGCGCATCCGC	60

Dh 1 ATGGGGCCCCGGCCAGCAICCTGCGGGGGCAAGCTGGACGCTGGGAGCGCATCCGC 60

Qy 61 CTGGCCCCGGGGCAAGAAGTGTACATGATGAAGACCTGCTGTGGCCAGCGCGAG 120

Dh 61 CTGGCCCCGGGGCAAGAAGTGTACATGATGAAGACCTGCTGTGGCCAGCGCGAG 120

Qy 121 CTGGAGAAGTTCGCCCTGAACCCGGCCCTGCTGAGAGACCGAGGGGCTGCAAGCAGAIC 180

Dh 121 CTGGAGAAGTTCGCCCTGAACCCGGCCCTGCTGAGAGACCGAGGGGCTGCAAGCAGAIC 180

Qy 181 ATCCGCCAGCTGCACCCCGCCCTGCAGACCGGGCAGGAGCTGAAGAGCCTGTCAAC 240

Dh 181 ATCCGCCAGCTGCACCCCGCCCTGCAGACCGGGCAGGAGCTGAAGAGCCTGTCAAC 240

Qy 241 ACCGTGGCCACCCCTGTACTGCTGACAGAGAAGATCGAGGTCGCGCACACCCAGAGGCC 300

Dh 241 ACCGTGGCCACCCCTGTACTGCTGACAGAGAAGATCGAGGTCGCGCACACCCAGAGGCC 300

Qy 301 CTGGACAAGATCGAGGAGGAGCAGAACAAAGTGCACGAGAAGATCCAGCAGGCGGAGGCC 360

Dh 301 CTGGACAAGATCGAGGAGGAGCAGAACAAAGTGCACGAGAAGATCCAGCAGGCGGAGGCC 360

Qy 361 GCCGACAAGGGCAAGGTGAGCCAGNACTACCCCATCGTGCAGAACCTGCAGGGCCAGATG 420

Dh 361 GCCGACAAGGGCAAGGTGAGCCAGNACTACCCCATCGTGCAGAACCTGCAGGGCCAGATG 420

Qy 421 GTGACACAGGCCATCAGCCCCCGCACCCCTGAACGCTGGTGAGGTGATCGAGGAGAG 480

Dh 421 GTGACACAGGCCATCAGCCCCCGCACCCCTGAACGCTGGTGAGGTGATCGAGGAGAG 480

Qy 481 GCCTTCAGCCCCGAGGTGATCCCCATGTTACCGGCCCTGAGCGAGGGGCCACCCGCCAG 540

Dh 481 GCCTTCAGCCCCGAGGTGATCCCCATGTTACCGGCCCTGAGCGAGGGGCCACCCGCCAG 540

Qy 541 GACCTGAACACGATGTTGAACACCGTGGCGGGCCACCGAGCGCCGATCGAGATGCTGAAG 600

Dh 541 GACCTGAACACGATGTTGAACACCGTGGCGGGCCACCGAGCGCCGATCGAGATGCTGAAG 600

Qy 601 GACACCATCAACGAGGAGGGCCGAGTGGACCGCGTGCACCCCGTGCACGGCGGCCCC 660

Dh 601 GACACCATCAACGAGGAGGGCCGAGTGGACCGCGTGCACCCCGTGCACGGCGGCCCC 660

Qy 661 ATCGCCCCCGCCAGATCGCGAGCCCCCGCGGAGGACATCGCGGGCACCCACCGACACC 720

Dh 661 ATCGCCCCCGCCAGATCGCGAGCCCCCGCGGAGGACATCGCGGGCACCCACCGACACC 720

Qy 721 CTGCAGGAGCAGATCGCTGGATGACCAAGAACCCCGGATCGCGGGTGGCGGACATCAIC 780

Dh 721 CTGCAGGAGCAGATCGCTGGATGACCAAGAACCCCGGATCGCGGGTGGCGGACATCAIC 780

Qy 781 AAGCGGTGGATCAICCTGGSCCTGAACAAGATCGTGGGATGTACAGCCCGGTGAGCATC 840

Dh 781 AAGCGGTGGATCAICCTGGSCCTGAACAAGATCGTGGGATGTACAGCCCGGTGAGCATC 840

Qy 841 CTGGACATCAAGCAGGGCCCCCAAGGAGCCCTTCCGGGACTACGTGGACCGCTTCTTCAAG 900

Dh 841 CTGGACATCAAGCAGGGCCCCCAAGGAGCCCTTCCGGGACTACGTGGACCGCTTCTTCAAG 900

Qy 901 ACCCTGGCGCCGAGCAGAGCACCCAGGAGGTGAAGAACTGGATGACCGACACCTGCTG 960

Dh 901 ACCCTGGCGCCGAGCAGAGCACCCAGGAGGTGAAGAACTGGATGACCGACACCTGCTG 960

Qy 961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCTCGGGGCTCTCGSCCGCGGGCGAGC 1020

Dh 961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCTCGGGGCTCTCGSCCGCGGGCGAGC 1020

Qy 1021 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGCGGGCCCCAGCCACAAGGCCGCTG 1080

Dh 1021 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGCGGGCCCCAGCCACAAGGCCGCTG 1080

Qy 1081 CTGGCCGAGGCGATGAGCCAGGCCAACACCGAGCGTGTATGACAGAGAGCAACITCAAG 1140

Dh 1081 CTGGCCGAGGCGATGAGCCAGGCCAACACCGAGCGTGTATGACAGAGAGCAACITCAAG 1140

Qy 1141 GGCCCCCGGGCAICGTCAAGTGTCTCACTGCGGCAAGGAGGGCCACATCGCCCCGAAC 1200

Dh 1141 GGCCCCCGGGCAICGTCAAGTGTCTCACTGCGGCAAGGAGGGCCACATCGCCCCGAAC 1200

Qy 1201 TGCCGGCCCCCGCCCAAGAGGGGCTGCTGGAAGTCCGCAAGGAGGGCCACCAAGATGAAG 1260

Dh 1201 TGCCGGCCCCCGCCCAAGAGGGGCTGCTGGAAGTCCGCAAGGAGGGCCACCAAGATGAAG 1260

Qy 1261 GACTGCACCGAGCGCCAGGCCAACTTCTGGGCAAGATCTGGCCCAAGCCACAAGGGCCGC 1320

Dh 1261 GACTGCACCGAGCGCCAGGCCAACTTCTGGGCAAGATCTGGCCCAAGCCACAAGGGCCGC 1320

Qy 1321 CCCGGCAACTTCTGAGAGCGCCCGGAGCCCAACCGCCCGCCCGCCCGGAGAGCTTCGC 1380

Dh 1321 CCCGGCAACTTCTGAGAGCGCCCGGAGCCCAACCGCCCGCCCGCCCGGAGAGCTTCGC 1380

Qy 1381 TTCGAGGAGACCAACCCCGCCAGAGCAGAGAGCAAGAGCCGCGAGACCTGACCAAGC 1440

Dh 1381 TTCGAGGAGACCAACCCCGCCAGAGCAGAGAGCAAGAGCCGCGAGACCTGACCAAGC 1440

Qy 1441 CTGAAGAGCCTGTTCTGGCAAGACCCCTGAGCCAGTAA 1479

Dh 1441 CTGAAGAGCCTGTTCTGGCAAGACCCCTGAGCCAGTAA 1479

RESULT 3
AX468547
LOCUS AX468547 1479 bp DNA linear PAT 16-JUL-2002
DEFINITION Sequence 67 from Patent WO0226209.
ACCESSION AX468547
VERSION AX468547.1 GI:21901377
KEYWORDS
SOURCE Human immunodeficiency virus 1 (HIV-1)
ORGANISM Human immunodeficiency virus 1
Viruses; Retroviral viruses; Retroviridae; Lentivirus; Primate
lentivirus group.
REFERENCE 1
AUTHORS O'Hagan,D., Otten,G., Donnelly,J.J., Polo,J.M., Barnett,S.,
Singh,M., Ulmer,J. and Dubensky,J.W.
TITLE Microparticles for delivery of the heterologous nucleic acids
JOURNAL Patent: WO 0226209-A 67 04-APR-2002;
CHIRON CORPORATION (US)
FEATURES
Location/Qualifiers
Source 1. 1479
/organism="Human immunodeficiency virus 1"
/mol_type="genomic DNA"
/db_xref="taxon:11676"

BASE COUNT 325 a 527 c 457 g 160 t 10 others
ORIGIN

Query Match 99.6%; Score 1473.4; DB 6; Length 1479;
Best Local Similarity 99.3%; Pred. No. 1.2e-150;
Matches 1458; Conservative 10; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGGGGCCCCGGCCAGCATCTGCGGGCGGCAAGCTGGACGCTGGGAGCGCATCCGC 60

Dh 1 ATGGGGCCCCGGCCAGCATCTGCGGGCGGCAAGCTGGACGCTGGGAGCGCATCCGC 60

Qy 61 CTGGCCCCCGGGCAAGAAGTGTACATGATGAAGACCTGCTGGAGCCGCGAG 120

Dh 61 CTGGCCCCCGGGCAAGAAGTGTACATGATGAAGACCTGCTGGAGCCGCGAG 120

Qy 121 CTGGAGAAGTTCGCCCTGAACCCCGCCCTGCTGGAGACCGAGGGCTGCAAGCAGATC 180

Dh 121 CTGGAGAAGTTCGCCCTGAACCCCGCCCTGCTGGAGACCGAGGGCTGCAAGCAGATC 180

Qy 181 ATCGGCCAGCTGCACCCCGCCCTGCAGACCGGGCAGGAGGAGCTGAAGAGCCTGTTCAAC 240

Dh 181 ATCGGCCAGCTGCACCCCGCCCTGCAGACCGGGCAGGAGGAGCTGAAGAGCCTGTTCAAC 240

Qy 241 ACCGTGGCCACCCCTGTACTGCTGACGAGAGAGATCGAGGTCCCGGACACCGAGGAGGCC 300

|||||

Db 241 ACCGTGGCCACCCCTGTACTGCGTGCCAGAGAGATCGAGGTSQCGCACACCAAGAGGCC 300

QY 301 CTGGACRAGATCGAGGAGGAGCAGACAAAGTGCACGAGAAGATCCAGCAGGCCAGGCC 360

Db 301 CTGGACRAGATCGAGGAGGAGCAGACAAAGTGCACGAGAAGATCCAGCAGGCCAGGCC 360

QY 361 GCCGACRAGGCAAGGTCAGCCAGAACTACCCCATCGTGCAGAACTCGCAGGGCCAGATG 420

Db 361 GCCGACRAGGCAAGGTCAGCCAGAACTACCCCATCGTGCAGAACTCGCAGGGCCAGATG 420

QY 421 GTGCACRAGGCCATCAGCCCCCGCACCCCTGAACGCCCTGGGTGAAGGTGATCGAGAGAG 480

Db 421 GTGCACRAGGCCATCAGCCCCCGCACCCCTGAACGCCCTGGGTGAAGGTGATCGAGAGAG 480

QY 481 GCCTTCAGCCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGGCCACCCCGAG 540

Db 481 GCCTTCAGCCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGGCCACCCCGAG 540

QY 541 GACCTGACACGATGTTGAACACCGTGGCGGCCACACCGCCCATCGAGATGCTCAAG 600

Db 541 GACCTGACACGATGTTGAACACCGTGGCGGCCACACCGCCCATCGAGATGCTCAAG 600

QY 601 GACACCATCAACGAGGAGGCCCGGAGTGGGACCGGTGCACCCCTGCACGCCCGGCC 660

Db 601 GACACCATCAACGAGGAGGCCCGGAGTGGGACCGGTGCACCCCTGCACGCCCGGCC 660

QY 661 ATCGCCCCCGCCAGATGCGGAGCGCCCGGCGAGGACATCGCCGGCACCAACGACAC 720

Db 661 ATCGCCCCCGCCAGATGCGGAGCGCCCGGCGAGGACATCGCCGGCACCAACGACAC 720

QY 721 CTGCAGGACGATCGCTGGATGACACGACCAACCCCCCATCCCGTGGGGGACATCTAC 780

Db 721 CTGCAGGACGATCGCTGGATGACACGACCAACCCCCCATCCCGTGGGGGACATCTAC 780

QY 781 AAGCGTGGATCATCTCGGCTGAACAAAGATCGTGCGGATGTACAGCCCGTGAGCATC 840

Db 781 AAGCCSTGGATCATCTCGGCTGAACAAAGATCGTGCGSNTGTACAGCCCGTGAGCATC 840

QY 841 CTGGACATCAAGCAGGGGCCCAAGGAGCCCTTCGCGACTACGTGGACCGCTTCCTCAAG 900

Db 841 CTGGACATCAAGCAGGGGCCCAAGGAGCCCTTCGCGACTACGTGGACCGCTTCCTCAAG 900

QY 901 ACCCTGCGCGCGAGCAGACACCCAGAGGTGAAGAACTGGATGACCGACACCCCTGCTG 960

Db 901 ACCCTGCGCGCGAGCAGACACCCAGAGGTGAAGAACTGGATGACCGACACCCCTGCTG 960

QY 961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCTGCGCGCTCTCGGCCCGCGCGAGC 1020

Db 961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCTGCGCGCTCTCGGCCCGCGCGAGC 1020

QY 1021 CTGGAGGAGATGATGACCGCTGCCAGGCGTGGGCGCCCGCCAGCCACAAAGCGCGCGTG 1080

Db 1021 CTGGAGGAGATGATGACCGCTGCCAGGCGTGGGCGCCCGCCAGCCACAAAGCGCGCGTG 1080

QY 1081 CTGGCCGAGCGGATGAGCCAGGCCAACACACGCGTGTATGATGCAGAGAGCAACTCAAG 1140

Db 1081 CTGGCCGAGCGSATGAGCCAGGCCAACACACGCGTGTATGATGCAGAGAGCAACTCAAG 1140

QY 1141 GGCCCCCGCGCATCTCAAGTGCCTCAACTGCGGCAAGGAGGGCCACATCGCCCGCAAC 1200

Db 1141 GGCCCCCGCGCATCTCAAGTGCCTCAACTGCGGCAAGGAGGGCCACATCGCCCGCAAC 1200

QY 1201 TGCCGCGCCCCCGCAAGAAGGGCTGTGGAAGTGCAGCAAGGAGGGCCACCAAGATGAAG 1260

Db 1201 TGCCGCGCCCCCGCAAGAAGGGCTGTGGAAGTGCAGCAAGGAGGGCCACCAAGATGAAG 1260

QY 1261 GACTGCACGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCAGCCAGCCAGGCGCG 1320

Db 1261 GACTGCACGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCAGCCAGCCAGGCGCG 1320

QY 1321 CCCGGCAACTTCCTGCAGAGCGCGCCAGGCCACCCCGCCCCCCCCCGCGGAGAGCTTCGC 1380

Db 1321 CCCGGCAACTTCCTGCAGAGCGCGCCAGGCCACCCCGCCCCCCCCCGCGGAGAGCTTCGC 1380

QY 1381 TTCGAGGAGACCCACCCCGCGCAGAAAGCAGAGAGCAAGCAGCCGCGAGACCCCTGACCAGC 1440

Db 1381 TTCGAGGAGACCCACCCCGCGCAGAAAGCAGAGAGCAAGCAGCCGCGAGACCCCTGACCAGC 1440

QY 1441 CTGAAGAGCCTGTTCTGGCAACGACCCCTGAGCCAGTAA 1479

Db 1441 CTGAAGAGCCTGTTCTGGCAACGACCCCTGAGCCAGTAA 1479

RESULT 4

AX455904

LOCUS AX455904

DEFINITION Sequence 20 from Patent WO204493.

ACCESSION AX455904

VERSION AX455904.1 GI:21714896

KEYWORDS

SOURCE synthetic construct

ORGANISM synthetic construct

artificial sequences.

REFERENCE 1

AUTHORS zur Megede,J., Barnett,S.W., Engselbrecht,S. and van Rensburg,E.

TITLE Polynucleotides encoding antigenic hiv type c polypeptides,

polypeptides and uses thereof

JOURNAL Patent: WO 0204493-A 20 17-JAN-2002;

CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)

FEATURES

Location/Qualifiers

1..1479

/organism="synthetic construct"

/mol_type="genomic DNA"

/db_xref="taxon:32630"

/note="synthetic gag coding sequence of HIV strain

AF110965"

BASE COUNT 325 a 533 c 461 g 160 t

ORIGIN

Query Match 98.9%; Score 1463; DB 6; Length 1479;

Best Local Similarity 99.3%; Pred. No. 1.6e-149;

Matches 1469; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 ATGGGGCGCCCGCCAGCATCTCTGCGGGCGGCAAGCTGAGACGCCTGGGAGCGCATCCGC 60

Db 1 ATGGGGCGCCCGCCAGCATCTCTGCGGGCGGCAAGCTGAGACGCCTGGGAGCGCATCCGC 60

QY 61 CTGCGCCCCCGGGCGCAAGAGTGTACATGATGAAGCACCTGGTGTGGGCCAGCCCGGAG 120

Db 61 CTGCGCCCCCGGGCGCAAGAGTGTACATGATGAAGCACCTGGTGTGGGCCAGCCCGGAG 120

QY 121 CTGGAGAAGTTCGCCCTGAACCCCGCCCTCTCTGGAGACCAAGGAGGCTGAAGAGCATC 180

Db 121 CTGGAGAAGTTCGCCCTGAACCCCGCCCTCTCTGGAGACCAAGGAGGCTGAAGAGCATC 180

QY 181 ATCCGCCAGCTGCACCCCGCCCTGCAGACCGCGCAGGAGGAGCTGAAGAGCCTGTTCAAC 240

Db 181 ATCCGCCAGCTGCACCCCGCCCTGCAGACCGCGCAGGAGGAGCTGAAGAGCCTGTTCAAC 240

QY 241 ACCGTGGCCACCCCTGTACTGCGTGCACGAGAAGATCGAGTCCCGGACACCAAGGAGGCC 300

Db 241 ACCGTGGCCACCCCTGTACTGCGTGCACGAGAAGATCGAGTCCCGGACACCAAGGAGGCC 300

QY 301 CTGGACAAGATCGAGGAGGAGGAGCAACAAGTCCAGCAGAGATCCAGAGGCCGAGGCC 360

Db 301 CTGGACAAGATCGAGGAGGAGGAGGAGCAACAAGTCCAGCAGAGATCCAGAGGCCGAGGCC 360

QY 361 GCCGACAAGGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACTTCGAGGGCCAGATG 420

Db 361 GCCGACAAGGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACTTCGAGGGCCAGATG 420

QY 421 GTGCACRAGGCCATCAGCCCCCGCACCCCTGAACGCCCTGGGTGAAGGTGATCGAGAGAG 480

Db 421 GTGCACRAGGCCATCAGCCCCCGCACCCCTGAACGCCCTGGGTGAAGGTGATCGAGAGAG 480

QY 481 GCCTTCAGCCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGGCCACCCCGAG 540

```

481 GCGTTCAGCCCGCAGGAGGATCCCGATGTTACCGCCCTGAGCGAGGGCGCCACGCCCGCAG 540
541 GACCTGAACACAGATGTTGAACACCGTGGGGGCGACACAGCGCGCATGCAGATGCTGAAG 600
541 GACCTGAACACAGATGTTGAACACCGTGGGGGCGACACAGCGCGCATGCAGATGCTGAAG 600
601 GACACCATCAACGAGGAGGCGCGCGAGTGGGAGCGCGTGCACCGCGTGCACCGCGCGCGCG 660
601 GACACCATCAACGAGGAGGCGCGCGAGTGGGAGCGCGTGCACCGCGTGCACCGCGCGCGCG 660
661 ATCGCCCGCGCGCGAGATGCGGAGCGCGCGCGGAGCGGACATGCGCGCGCGCGCGCGCGCG 720
661 ATCGCCCGCGCGCGAGATGCGGAGCGCGCGCGGAGCGGACATGCGCGCGCGCGCGCGCGCG 720
721 CTGAGGAGCAGATGCGCGCGCGCGCGCGGAGCGGACATGCGCGCGCGCGCGCGCGCGCG 780
721 CTGAGGAGCAGATGCGCGCGCGCGCGCGGAGCGGACATGCGCGCGCGCGCGCGCGCGCG 780
781 AAGCGGTGATCATCTCGCGCGCGCGCGGAGCGGACATGCGCGCGCGCGCGCGCGCGCGCG 840
781 AAGCGGTGATCATCTCGCGCGCGCGCGGAGCGGACATGCGCGCGCGCGCGCGCGCGCGCG 840
841 CTGACATCAAGCAGGCGCGCGCGGAGCGGACATGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
841 CTGACATCAAGCAGGCGCGCGCGGAGCGGACATGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
901 ACCCTGCGCGCGCGCGCGGAGCGGACATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 960
901 ACCCTGCGCGCGCGCGGAGCGGACATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 960
961 GTGAGAACCGCGCGCGCGCGGAGCGGACATGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020
961 GTGAGAACCGCGCGCGCGGAGCGGACATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020
1021 CTGAGGAGATGATGACCGCGCGCGCGGAGCGGACATGCGCGCGCGCGCGCGCGCGCGCG 1080
1021 CTGAGGAGATGATGACCGCGCGCGCGGAGCGGACATGCGCGCGCGCGCGCGCGCGCGCG 1080
1081 CTGCGCGAGGCGCGCGCGCGGAGCGGACATGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1140
1081 CTGCGCGAGGCGCGCGCGGAGCGGACATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1140
1141 GCGCGCGCGCGCGCGCGGAGCGGACATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1200
1141 GCGCGCGCGCGCGCGGAGCGGACATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1200
1201 TCGCGCGCGCGCGCGCGGAGCGGACATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1260
1201 TCGCGCGCGCGCGCGGAGCGGACATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1260
1261 GACTGCGCGCGCGCGCGCGGAGCGGACATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1320
1261 GACTGCGCGCGCGCGCGGAGCGGACATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1320
1321 CCGCGCGCGCGCGCGCGGAGCGGACATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
1321 CCGCGCGCGCGCGCGGAGCGGACATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
1381 TTCGAGGAGACCGCGCGCGCGGAGCGGACATGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1440
1381 TTCGAGGAGACCGCGCGCGCGGAGCGGACATGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1440
1441 CTGAGAGCGCTGTTGCGGCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1479
1441 CTGAGAGCGCTGTTGCGGCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1479
```

RESULT 5
AX455983
LOCUS
DEFINITION Sequence 99 from Patent WO0204493.
ACCESSION AX455983

```

VERSION AX455983.1 GI:21714967
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.
TITLE Polynucleotides encoding antigenic hiv type c polypeptides,
JOURNAL polypeptides and uses thereof
PATENT Patent: WO 0204493-A 99 17-JAN-2002;
CHIRON CORPORATION (US); University of Stellenbosch (ZA)
FEATURES
Location/Qualifiers
source 1..1491
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="Gag_TV2_C_2hopt"
BASE COUNT 319 a 557 c 457 g 158 t
ORIGIN
Query Match 87.1%; Score 1288.8; DB 6; Length 1491;
Best Local Similarity 93.0%; Pred. No. 1.1e-140;
Matches 1389; Conservative 0; Mismatches 87; Indels 18; Gaps 3;
QY 1 ATGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
DB 1 ATGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
QY 61 CTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
DB 61 CTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
QY 121 CTGGAGAGCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
DB 121 CTGGAGAGCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
QY 181 ATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
DB 181 ATCAAGCAGCTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
QY 241 ACCGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
DB 241 ACCGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
QY 301 CTGGACACAGATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
DB 301 CTGGACACAGATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
QY 361 GCGGACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420
DB 361 GCGGACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420
QY 421 GTGCACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
DB 421 GTGCACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
QY 477 GTGCACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
DB 477 GTGCACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
QY 541 GCGGACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
DB 541 GCGGACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
QY 597 GACCTGAACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 657
DB 597 GACCTGAACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 657
QY 661 ATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
DB 661 ATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
QY 717 GTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780
DB 721 CTGCAGGAGCAGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780
```

```

|||||
718 CTGCAGGACGATCGCTGGATGACCAAGCAACCCCGCATCCCGTGGCGACATCTAC 777
|||||
781 AAGCGGTGGATCATCTGGGCTGAAACAAGATCGTGGGATGTACAGCCCGGTGAGCATC 840
|||||
778 AAGCGGTGGATCATCTGGGCTGAAACAAGATCGTGGGATGTACAGCCCGGTGAGCATC 837
|||||
841 CTGGACATCAAGCAGGCGCCCAAGAGCCCTTCCGGGACTACGTGAGCCGCTTCTTCAAG 900
|||||
838 CTGGACATCAAGCAGGCGCCCAAGAGCCCTTCCGGGACTACGTGAGCCGCTTCTTCAAG 897
|||||
901 ACCCTGGCGCGGAGCAGACACCCAGGAGGTGAAGAACTGGATGACCGACCCCTGCTG 960
|||||
898 ACCCTGGCGCGGAGCAGACACCCAGGAGGTGAAGAACTGGATGACCGACACCCCTGCTG 957
|||||
961 GTGCAGAACGCCAACCCGACTGCAAGACCATCTGCGCGCTCTCGGCCCGCGGCCAGC 1020
|||||
958 GTGCAGAACGCCAACCCGACTGCAAGACCATCTGCGCGCTCTCGGCCCGCGGCCAGC 1017
|||||
1021 CTGGAGGAGATGATGACCGCTGCGAGGCGTGGGCGGCCCGAGCAAGGCCCGCGTG 1080
|||||
1018 CTGGAGGAGATGATGACCGCTGCGAGGCGTGGGCGGCCCGAGCAAGGCCCGCGTG 1077
|||||
1081 CTGGCGGAGCGGATGACCCAGGC---CAACACCAAGCGCTGATGATGCAAGAGCAACTTC 1137
|||||
1078 CTGGCGGAGCGGATGACCCAGGCCAACACACCAAGCGTGTGATGATGATGATGATGATG 1137
|||||
1138 AAGGCGCCCGCGGCGATGCTCAAGTGTCTCAACTGCGCAAGGAGGCGCCACATCGCCCGC 1197
|||||
1138 AAGGCGCCCGCGGCGGCTGAAGTGTCTCAACTGCGCGCGCGAGGCGCCACATCGCCCGC 1197
|||||
1198 AACTGCGCGCGCGCGCGCAAGAGGGCTGCTGGAAGTGGCGCAAGAGGGCGCCACAGATG 1257
|||||
1198 AACTGCGCGCGCGCGCGCAAGCGCGGCTGCTGGAAGTGGCGCAAGAGGGCGCCACAGATG 1257
|||||
1258 AAGGACTGCACCGAGCGCCAGGCCAACTTCTGGGCAAGATCTGGGCGCGCCACAGAGGGC 1317
|||||
1258 AAGGACTGCACCGAGCGCCAGGCCAACTTCTGGGCAAGATCTGGGCGCGCCACAGAGGGC 1317
|||||
1318 CGCCCGCGCAACTTCTGAGAGCGCGCGCGAGCGCCAGCGCGCGCGCGCGCGCGCGCGCG 1377
|||||
1318 CGCCCGCGCAACTTCTGAGAGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1377
|||||
1378 CGCTTCGAGG-----AGACCAACCCCGCGCGCAAGAGCGAGGAGCAAGAGCGCG 1425
|||||
1378 GCGCCCGCGCGCGGAGAGGTTCAAGTTCAAGGAGACCGCGCAAGAGCGCGCGCGCGCGCG 1437
|||||
1426 GAGACCTGACCGAGCGCTGGAAGCGCTTCTGGGCAAGAGCGCGCGCGCGCGCGCGCGCG 1479
|||||
1438 GAGCCCTGACCGAGCGCTGGAAGCGCTTCTGGGCAAGAGCGCGCGCGCGCGCGCGCGCG 1491
|||||
```

```

RESULT 6
AX455935
LOCUS AX455935 1494 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 51 from Patent WO0204493.
ACCESSION AX455935
VERSION AX455935.1 GI:21714919
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE
1
AUTHORS zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.
TITLE Polynucleotides encoding antigenic hiv type c polypeptides,
polypeptides and uses thereof
JOURNAL Patent: WO 0204493-A 51 17-JAN-2002;
CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)
FEATURES
source
1. .1494
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
```

```

/note="HIV Type C Gag optimized"
BASE COUNT 325 a 557 c 455 g 157 t
ORIGIN
Query Match 87.1%; Score 1288; DB 6; Length 1494;
Best Local Similarity 92.9%; Pred. No. 1.3e-130;
Matches 1393; Conservative 0; Mismatches 80; Indels 27; Gaps 3;
QY 1 ATGGGGCGCGCGCGCAGCATCTCTGCGCGCGCGCAAGCTGGAGCGCTGGGAGCGCATCCGC 60
Db 1 ATGGGGCGCGCGCGCAGCATCTCTGAGCGCGCGCAAGCTGGAGCGCATCCGC 60
QY 61 CTGGCGCGCGCGCGCAAGAGTGTCTACATGATGAAGCACCTGGTGTGGCGCCAGCGCGAG 120
Db 61 CTGGCGCGCGCGCGCAAGAGTGTCTGCTGAAGCACCTGGTGTGGCGCCAGCGCGAG 120
QY 121 CTGGAGAGTTCGCCCTGACCCCGCGCTGCTGGAGACCGAGGGCTGCAAGCAGATC 180
Db 121 CTGGAGCGCTTCGCCCTGACCCCGCGCTGCTGGAGACCGAGGGCTGCAAGCAGATC 180
QY 181 ATCCGCGCAGCTGCACCCCGCTGCAGACCGCGGAGGAGCTGAAGAGCGCTGTTCAC 240
Db 181 ATCAAGCAGCTGCAGCCCGCTGCAGACCGCGGACCGAGGAGCTGCGCAGCGCTGTTCAC 240
QY 241 ACCGTGGCCACCCCTGTACTGCTGTCACGAGAAGATCGAGTCCCGGACACCAAGGAGGC 300
Db 241 ACCGTGGCCACCCCTGTACTGCTGTCACAAAGGCATCGAGTTCGCGACACCAAGGAGGC 300
QY 301 CTGGACAAGATCGAGGAGGAGCAACAAGTGCAGACGAGGAGTCCAGAGGCGCGAGGC 360
Db 301 CTGGACAAGATCGAGGAGGAGCAACAAGTGCAGACGAGGCGCGAGGCGCGCAAGGC 360
QY 361 GCCGACAAGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACTCGAGGCGCGAGATG 420
Db 361 GCCGACGAG---AAGGTGAGCAAACTACCCCATCGTGCAGAACTCGAGGCGCGAGATG 417
QY 421 GTGCACCGAGCGCATCAGCCCGCGCACCCCTGAACGCGCTGNTGAAGGTGATCGAGGAGAG 480
Db 418 GTGCACCGAGCGCATCAGCCCGCGCACCCCTGAACGCGCTGATCAAGGTGATCGAGGAGAG 477
QY 481 GCCTTCAGCCCGGAGGTGATCCCCATGTTTACCGCGCGGAGGCGCGCGCGCGCGCGCG 540
Db 478 GCCTTCAACCCCGGAGGATCCCCATGTTTACCGCGCGGAGGCGCGCGCGCGCGCGCG 537
QY 541 GACCTGAACACGATGTTGACACCGCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
Db 538 GACCTGAACACGATGTTGACACCGCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 597
QY 601 GACACCATCAAGCAGGAGGCGCGCGGAGTGGGAGCGCGGTGACCCCGTGCAGCGCGCGCG 660
Db 598 GACACCATCAAGCAGGAGGCGCGCGGAGTGGGAGCGCGCGCGCGCGCGCGCGCGCGCG 657
QY 661 ATCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
Db 658 GTGGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 717
QY 721 CTGCAGGAGCAGATCGCTGGATGACCGAGCAACCCCGCGCGCGCGCGCGCGCGCGCGCG 780
Db 718 CTGCAGGAGCAGATCGCTGGATGACCGAGCAACCCCGCGCGCGCGCGCGCGCGCGCG 777
QY 781 AAGCGGTGGATCATCTGGGCGCTGAACAAGATCGTGGGATGTACAGCCCGCGTGGAGCATC 840
Db 778 AAGCGGTGGATCATCTGGGCGCTGAACAAGATCGTGGGATGTACAGCCCGCGTGGAGCATC 837
QY 841 CTGGACATCAAGCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
Db 838 CTGGACATCAAGCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 897
QY 901 ACCCTGCGCGCGAGCAGACACCCAGGAGGTGAAGAACTGGATGAGCGCGCTTCTTCAAG 960
Db 898 ACCCTGCGCGCGAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 957
QY 961 GTGCAGAACGCCAACCCGACTGCAAGACCATCTGCGCGCTCTCGGCCCGCGGCCAGC 1020
```


1198 AACTGCGCGCCCGCGAAGAGGGCTGCTGGAGTGGCGCAAGGAGGGCCACAGATG 1257
1258 AAGGACTGCACCGGAGCGCCAGCCAACTTCTTGGCAAGATCTGGCCAGCCACAGGGC 1317
1258 AAGGACTGCACCGGAGCGCCAGCCAACTTCTTGGCAAGATCTGGCCAGCCACAGGGC 1317
1318 CGCCCGCGCAACTTCTTGCAGAGCCG-----CGCCGAGCCACCGCC 1359
1318 CGCCCGCGCAACTTCTTGCAGAACCGCAGCGAGCCCGCCCGCCCGTGTCCACCGCC 1377
1360 CCGCCCGCGGAGAGCTTCCGCTTCGAGGAGACCCCGCCCGCCAGAGAGGAGGAG 1419
1378 CCGCCCGCGGAGAGCTTCCGCTTCGAGGAGACCCCGCCCGCCCGCAGGAGGAGG 1437
1420 G-----ACCGCGAGACCTGACAGCCCTGAGAGCCCTGTTGGGCAACGAGCC 1467
1438 GACCGGAGCCCTACCGGAGCCCTGACCGCCCTGCGAGCCCTGTTGGGAGCGGCC 1497
1468 CTGAGCCAGTAA 1479
1498 CTGAGCCAGTAA 1509
RESULT 8
AX468544
LOCUS AX468544 1509 bp DNA linear PAT 16-JUL-2002
DEFINITION Sequence 64 from Patent WO0226209.
ACCESSION AX468544
VERSION AX468544.1 GI:21901374
KEYWORDS
SOURCE Human immunodeficiency virus 1 (HIV-1);
ORGANISM Human immunodeficiency virus 1
Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
lentivirus group.
REFERENCE 1
AUTHORS O'Hagan,D., Otten,G., Donnelly,J.J., Polo,J.M., Barnett,S.,
Singh,M., Uimer,J. and Dubensky,T.W.
TITLE Microparticles for delivery of the heterologous nucleic acids
JOURNAL Patent: WO 0226209-A 64 04-APR-2002;
CHIRON CORPORATION (US)
FEATURES
Location/Qualifiers
Source 1..1509
/organism="Human immunodeficiency virus 1"
/mol_type="genomic DNA"
/db_xref="taxon:11676"
BASE COUNT 320 a 555 c 472 g 162 t
ORIGIN
Query Match 96.2%; Score 1275.2; DB #; Length 1509;
Best Local Similarity 92.5%; Pred. No. 3.2e-129;
Matches 1398; Conservative 0; Mismatches 78; Indels 36; Gaps 4;
1 ATGGGCGCGCGCGCCAGCATCTCTGGCGGCGCAAGCTGGAGCCCTGGGAGCCGATCCGC 60
1 ATGGGCGCGCGCGCGCCAGCATCTCTGGCGGCGCAAGCTGGAGCAAGTGGGAGATCCGC 60
61 CTGGCGCGCGCGCGCGCAAGAGTGTCTATGATGAAGCAACCTGGTGTGGCGAGCGCGAG 120
61 CTGGCGCGCGCGCGCGCAAGAGTGTCTATGATGAAGCAACCTGGTGTGGCGAGCGGTGAG 120
121 CTGGAGAGTTCGCGCTGAACCGCGCGCTGCTGGAGACCGAGGAGGCTGAAGAGCAGATC 180
121 CTGGAGGCTTCGCGCTGAACCGCGCGCTGCTGGAGACCGCGAGGAGGCTGAAGAGCAGATC 180
181 ATCCCGCAGCTGCACCCCGCCCTGCAGACCGCGAGCGAGGAGTGAAGAGCAGATC 240
181 ATGAGAGAGCTGCAGCCCGCGCTGCAGACCGCGAGCGAGGAGTGGCGAGCGCTGTACAAC 240
241 ACCGTGGCCACCCCTGTACTGGTGCACGAGAGATCGAGGTCCGCGACACCCAGGAGGCC 300
241 ACCGTGGCCACCCCTGTACTGGTGCACGCGCGGATCGAGGTCCGCGACACCCAGGAGGCC 300

301 CTGGACAAAGATCCAGGAGGAGGAGCAACAAGTGGCAGACAGAGATCCAGCGCCGAGGCC 360
301 CTGGACAAAGATCCAGGAGGAGGAGGAGCAACAAGTGGCAGACAGAGATCCAGCGCCGAGGAG 360
361 GCGGACAAAGGCAAGGTGAGCCAGAACTACCCCATCTGTCAGAACTCTGCAGGGCCAGATG 420
361 GCGGAC--GGCAAGGTGAGCCAGAACTACCCCATCTGTCAGAACTCTGCAGGGCCAGATG 417
421 GTGCACCAAGGCGCATCAGCCCGCCAGCCCTGAACGCTGGTGAAGGTGATCGAGGAGAAG 480
418 GTGCACCAAGGCGCATCAGCCCGCCAGCCCTGAACGCTGGTGAAGGTGATCGAGGAGAAG 477
481 GCCTTCAGCCCGGAGGTGATCCCATCTGTCAGCCCGCTGAGGAGGGCGCCACCCCGCAG 540
478 GCCTTCAGCCCGGAGGTGATCCCATCTGTCAGCCCGCTGAGGAGGGCGCCACCCCGCAG 537
541 GACCTGAACACCGATGTTGAACACCGTGGCGGCGCACCAAGCGCCCATGCAGATGCTGAAG 600
538 GACCTGAACACCGATGTTGAACACCGTGGCGGCGCACCAAGCGCCCATGCAGATGCTGAAG 597
601 GACACCATCAACGAGGAGGCGCGCGAGTGGGAGCGCGTGCACCCCGTGCACGCGCGGCC 660
598 GACACCATCAACGAGGAGGCGCGCGAGTGGGAGCGCGTGCACCCCGTGCAGGCGCGGCC 657
661 ATCGCCCGCGCGCATGCGGAGCGCGCGCGGAGCGAGCATCGCGGCGCACCAAGCGCACCC 720
658 GTGSCCGCGCGCGCATGCGGAGCGCGCGCGCGGAGCGAGCATCGCGGCGCGCACCAAGCGCACCC 717
721 CTGCAGGAGCAGATCGCCCTGGATGACCAAGCAACCCCGCATCGCGCTGGCGGACATCTAC 780
718 CTGCAGGAGCAGATCGCCCTGGATGACCAAGCAACCCCGCATCGCGCTGGCGGACATCTAC 777
781 AAGCGTGGATCATCTCTGGCGCTGAACAAGATCTGGCGGATGATACAGCCCGCTGAGCATC 840
778 AAGCGTGGATCATCTCTGGCGCTGAACAAGATCTGGCGGATGATACAGCCCGCTGAGCATC 837
841 CTGGACATCAAGCAGGCGCGCGCGCGCGCGGAGCGCTTCCCGGACTACGTGACCGCGCTTCTCAAG 900
838 CTGGACATCG 897
901 ACCCTGCG 960
898 ACCCTGCG 957
961 GTGCAGAAACG 1020
958 GTGCAGAAACG 1017
1021 CTGGAGGAGATGATGCG 1080
1018 CTGGAGGAGATGATGCG 1077
1081 CTGGCGGAGGCGATGAGCCAGGCGCAACACCG--AGCGTCATGATGCAGAAAGAGCAACTTC 1137
1078 CTGGCGGAGGCGATGAGCCAGGCGCAACAGCGGTGAACATCATGATGCAGAAAGAGCAACTTC 1137
1138 AAGGCG 1197
1138 AAGGCG 1197
1198 AACTGCG 1257
1198 AACTGCG 1257
1258 AAGGACTGCACCGGAGCG 1317
1258 AAGGACTGCACCGGAGCG 1317
1318 CGCCCGCGCGCAACTTCTTGCAGAGCGCG-----CGCCGAGCCACCGCC 1359
1318 CGCCCGCGCGCAACTTCTTGCAGAGCG 1377
1360 CCGCCCGCGGAGAGCTTCCGCTTCGAGGAGACCCCGCCCGCCAGAGAGGAGGAGGAG 1419
1378 CCGCCCGCGGAGAGCTTCCGCTTCGAGGAGACCCCGCCCGCCCGCAGGAGGAGGAGG 1437
1420 G-----ACCGCGAGACCTGACAGCCCTGAGAGCCCTGTTGGGCAACGAGCC 1467
1438 GACCGGAGCCCTACCGGAGCCCTGACCGCCCTGCGAGCCCTGTTGGGAGCGGCC 1497
1468 CTGAGCCAGTAA 1479
1498 CTGAGCCAGTAA 1509

Db 1378 CCCCCCGGAGAGCTTCGCTTCGAGGAGACCCACCCCGCCCGCCAGCAGGAGGCGCAAG 1437
QY 1420 G-----ACCGGAGACCCCTGACCAGCCCTGAGAGCCCTGTTGCGCAACACCC 1467
Db 1438 GACCGGAGCCCTACCGCGAGCCCTGACCGCCCTGCGCAGCCCTGCGGAGGCGGCCCC 1497
QY 1468 CTGAGCCAGTAA 1479
Db 1498 CTGAGCCAGTAA 1509

RESULT 9
AX468548
LOCUS AX468548
DEFINITION Sequence 68 from Patent: WO0226209.
ACCESSION AX468548
VERSION AX468548.1 GI:21901378
KEYWORDS
SOURCE Human immunodeficiency virus 1 (HIV-1)
ORGANISM Human immunodeficiency virus 1
Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
lentivirus group.
REFERENCE 1
AUTHORS O'Hagan,D., Otten,G., Donnelly,J.J., Pollo,J.M., Barnett,S.,
Singh,M., Ulmer,J. and Dubensky,T.W.
TITLE Microparticles for delivery of the heterologous nucleic acids
JOURNAL Patent: WO 0226209-A 68 04-APR-2002;
CHIRON CORPORATION (US)
FEATURES
source Location/Qualifiers
1..1509 /organism="Human immunodeficiency virus 1"
/mol_type="genomic DNA"
/db_xref="taxon:11676"
BASE COUNT 320 a 552 c 467 g 158 t 12 others
ORIGIN

Query Match 86.1%; Score 1273.6; DB 6; Length 1509;
Best Local Similarity 91.8%; Pred. No. 4.8e-129;
Matches 1388; Conservative 12; Mismatches 76; Indels 36; Gaps 4:

QY 1 ATGGGCGCGCGCCAGCATCTCTGCGGCGCGCAAGTGGACGCTGGGAGCGCATCCCG 60
Db 1 ATGGGCGCGCGCCAGCATCTCTGCGGCGCGCAAGTGGACGCTGGGAGAGATCCCG 60

QY 61 CTGCGCCCGCGCGGCAAGTGTCTACATGATGAAGACCTGGTGTGGCCCGCAGCGAG 120
Db 61 CTGCGCCCGCGCGGCAAGTGTCTACATGATGAAGACCTGGTGTGGCCCGCAGCGAG 120

QY 121 CTGGAGAGTTCGCCCTGAACCCCGGCTCTGGAGACCGAGGCTGCAAGCAGATC 180
Db 121 CTGGAGGCTTCGCCCTGAACCCCGGCTCTGGAGACCGGCTGCAAGCAGATC 180

QY 181 ATCCGCCAGCTGCACCCCGCCCTGCAGACCGGCGAGCGGAGCTGAGAGCCCTGTCAAC 240
Db 181 ATGAAGCAGCTGCAGCCCGCCCTGCAGACCGGCGAGCGGAGCTGCGCAGCCCTGTACAC 240

QY 241 ACCGTGGCCAGCCCTGTACTGCTGCACGAGAAGATCGAGGTCCGCGACACCAAGAGGCC 300
Db 241 ACCGTGGCCAGCCCTGTACTGCTGCACCCCGGCTCGAGGTSCGCGACACCAAGAGGCC 300

QY 301 CTGGACAAGATCGAGGAGCAGAGCAAGTSCCAGCAGAAGATCCAGCAGGCGCGAGGCC 360
Db 301 CTGGACAAGATCGAGGAGCAGAGCAAGWSCCAGCAGAAGACCCAGCAGGCCAAGAG 360

QY 361 GCCGACAAGGCAAGGTGAGCCAGAACTACCCCATCTGCAGAACTGCAGGGCCAGATG 420
Db 361 GCCGAC---GGCAAGGTGAGCCAGAACTACCCCATCTGCAGAACTGCAGGGCCAGATG 420

QY 421 GTGCACCAAGGCCATCAGCCCCCGCACCTGAAACGCTGGGTGAAGGTGATCGAGAGAG 480
Db 418 GTGCACCAAGGCCATCAGCCCCCGCACCTGAAACGCTGGGTGAAGGTGATCGAGAGAG 477

QY 481 GCCTTCAGCCCCGAGGTGATCCCCAIGTCAACCGCCCTAGCGAGGGCGCCACCCCCAG 540
Db 478 GCCTTCAGCCCCGAGGTGATCCCCATGTCACCGCCCTGAGCGAGGGCGCCACCCCCAG 537

QY 541 GACCTGAACACAGTAITTTGAACACCGTGGGGGCCACCCAGGCCGCCATGAGATGCTGAAG 600
Db 538 GACCTGAACACASAITGTGAACACCGTGGGGGCCACCCAGGCCGCCATGAGATGCTGAAG 597

QY 601 GACACCATCAACGAGGAGCGCGGAGTGGGAGCGCGTGCACCCCGTGCAGCGCGGCC 660
Db 598 GACACCATCAACGAGGAGCGCGGAGTGGGAGCGCGTGCACCCCGTGCAGCGCGGCC 657

QY 661 ATCGCCCCCGCCAGATGCGCGAGCCCCCGCGGAGCGACATCGCGGACACCAAGCACC 720
Db 658 GTGGCCCCCGCCAGATGCGCGAGCCCCCGCGGAGCGACATCGCGGCGCCACCAAGCACC 717

QY 721 CTGCAGGAGCAGATCGCCTGGATGACCAAGACCCCGCCATCCCGTGGGCGACATCTAC 780
Db 718 CTGCAGGAGCAGATCGCCTGGATGACCAAGACCCCGCCATCCCGTGGGCGACATCTAC 777

QY 781 AAGCGTGGATCATCTCTGGGCTGAACAAAGATCGTGGGATGTACAGCCCGTGGAGATC 840
Db 778 AAGCGTGGATCATCTCTGGGCTGAACAAAGATCGTGGGATGTACAGCCCGTGGAGATC 837

QY 841 CTGGACATCAACGAGGCGCCCAAGGACCTTCGCGACTACGTTGACCGCTTCTTCAAG 900
Db 838 CTGGACATCCCGCAGGCGCCCAAGGACCTTCGCGACTACGTTGACCGCTTCTTCAAG 897

QY 901 ACCCTGCGCGCGGAGCAGACACCCAGGAGTGAAGAATCGATGACCAACCCCTGCTG 960
Db 898 ACCCTGCGCGCGGAGCAGACACCCAGGAGTGAAGAATCGATGACCAACCCCTGCTG 957

QY 961 CTGCAGAACGCCAACCCCGACTGCAAGACCATCTCTGCGGCTCTCGGCCCGCGCGCAGC 1020
Db 958 GTGCAGAACGCCAACCCCGACTGCAAGACCATCTCTGCGGCTCTCGGCCCGCGCGCAGC 1017

QY 1021 CTGGAGGAGATGATGACCGCTGCCAGGCGCTGGGCGGCGGCGGCGGCGGCGGCGTG 1080
Db 1018 CTGGAGGAGATGATGACCGCTGCCAGGCGCTGGGCGGCGGCGGCGGCGGCGGCGTG 1077

QY 1081 CTGGCCGAGGCGATGAGCCAGGCGCAACACC---AGCGTATGATGCAAGAGCAACTTC 1137
Db 1078 CTGGCCGAGGCGATGAGCCAGGCGCAACAGCGTGAACATCATGATGCAAGAGCAACTTC 1137

QY 1138 AAGGGCCCCCGCGGCATCTCAAGTGTCTCAACTGCGGCAAGGAGGCGGCGGCGGCG 1197
Db 1138 AAGGGCCCCCGCGCAACGTSAACTGCTTCAACTGCGGCAAGGAGGCGGCGGCGGCG 1197

QY 1198 AACTGCCGCGCGCCCGCCAGCAAGGCTGCTGGAAGTGGCGCAAGGAGGCGGCGGCGG 1257
Db 1198 AACTGCCGCGCGCCCGCCAGCAAGGCTGCTGGAAGTGGCGCAAGGAGGCGGCGGCGG 1257

QY 1258 AAGGACTGCACCGAGCGCGGCGGCGCAACTTCTCTGGCAAGATCTGGCCCGGCGGCGG 1317
Db 1258 AAGGACTGCACCGAGCGCGGCGGCGCAACTTCTCTGGCAAGATCTGGCCCGGCGGCGG 1317

QY 1318 CGCCCCGCGCACTTCTCTGCAAGCGG-----CCCCGAGCCCCACCGCC 1359
Db 1318 CGCCCCGCGCACTTCTCTGCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1377

QY 1360 CCCCCCGCGGAGAGCTTCCGCTTCGAGGAGACCAACCCCGCGGCGGCGGAGGAGGCGAG 1419
Db 1378 CCCCCCGCGGAGAGCTTCCGCTTCGAGGAGACCAACCCCGCGGCGGCGGCGGCGGCGG 1437

QY 1420 G-----ACCGGAGACCCCTGACCAAGCCTGAAGAGCCTGTTCGGCAACGACCCC 1467
Db 1438 GACCGGAGCCCTACCGCGGAGCCCTGACCGCCCTTGCAGCCCTTGCAGCAGCCTGTTC 1497

QY 1468 CTGAGCCAGTAA 1479
Db 1498 CTGAGCCAGTAA 1509

Matches 1344; Conservative 0; Mismatches 130; Indels 20; Gaps 4;			
QY	1	ATGGGCGCGCGCCAGCATCTCTGGCGGCGGCAAGCTGGACGCCCTGGGAGCGCATCCGC	60
Db	1880	ATGGGCGCGCGCCAGCATCTCTGGCGGCGGCAAGCTGGACGCCCTGGGAGCGCATCAAG	1939
QY	61	CTGGCGCCCGCGGCAAGAGTGTCTACATGATGAAGACACCTGGTGGGCGCCAGCCGAG	120
Db	1940	CTGAAGCCCGCGGCAAGAGCACTACATGATGAAGACACCTGGTGGGCGCCAGCCGAG	1999
QY	121	CTGGAGAGTTCCGCTGAACCCCGGCTGCTGGAGACCGAGGAGGCTGCAAGAGATC	180
Db	2000	CTGGAGCGCTTCGCTGACCCCGGCTGCTGGAGACCGAGGAGGCTGCAAGAGATC	2059
QY	181	ATCCGCGCAGTGCACCCCGCTGACAGCCGCGCAGCGGAGGAGGCTGCAAGAGATC	240
Db	2060	ATGAAGCAGTGCAGCCCGCTGACAGCCGCGCAGCGGAGGAGGCTGCAAGAGATC	2119
QY	241	ACCGTGGCCACCTGTACTCGGTGCACGAGAAGATCGAGGTCCGCGACACCAAGGAGGC	300
Db	2120	ACCGTGGCCACCTGTACTCGGTGCACGAGAAGATCGAGGTCCGCGACACCAAGGAGGC	2179
QY	301	CTGGACAGATCGAGGAGCAGAACAGTGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGG	360
Db	2180	CTGGACAGATCAAGGAGGAGCAGAACAGTGCAGAGGAGGAGGAGGAGGAGGAGGAGG	2239
QY	361	GCCGACAAAGGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	420
Db	2240	GCCGACAAAGGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	2299
QY	421	GTGCACAGCCCATCAGCCCGCGCAGCCCTGAACGCGCTGGTGAAGGTGATCGAGGAGAA	480
Db	2300	GTGCACAGCCCATCAGCCCGCGCAGCCCTGAACGCGCTGGTGAAGGTGATCGAGGAGAA	2359
QY	481	GCCTTCAGCCCGGAGGTGATCCCATGTTACACCGCCCTGAGCGAGGCGCCACCCGAG	540
Db	2360	GCCTTCAGCCCGGAGGTGATCCCATGTTACACCGCCCTGAGCGAGGCGCCACCCGAG	2419
QY	541	GACCTGAACAGATGTTGAACACCGTGGCGGCGCCACCGAGCGCCGATGAGATGCTGAAG	600
Db	2420	GACCTGAACAGATGCTGAACACCGTGGCGGCGCCACCGAGCGCCGATGAGATGCTGAAG	2479
QY	601	GACACCATCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	660
Db	2480	GACACCATCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	2539
QY	661	ATCGCCCGCGCCAGATGCGGAGCGCCGCGGAGCGGAGGAGGAGGAGGAGGAGGAGG	720
Db	2540	ATCGCCCGCGCCAGATGCGGAGCGCCGCGGAGCGGAGGAGGAGGAGGAGGAGGAGG	2599
QY	721	CTGCAGGAGCAGATCGCTGGATGACCAAGACCCCGCCGATCGCCGCGGAGCATCTAC	780
Db	2600	CTGCAGGAGCAGATCACCCTGATGACCAAGACCCCGCCGATCGCCGCGGAGCATCTAC	2659
QY	781	AAGCGGTGATCATCTGGGCTGAACAGATCGTGGGATGTACAGCCCGCGTGAAGATC	840
Db	2660	AAGCGGTGATCATCTGGGCTGAACAGATCGTGGGATGTACAGCCCGCGTGAAGATC	2719
QY	841	CTGGACATCAAGCAGGCGCCCAAGGAGGCTTCCGCGAGTACGTGGACCGCTTCTCAAG	900
Db	2720	CTGGACATCAAGCAGGCGCCCAAGGAGGCTTCCGCGAGTACGTGGACCGCTTCTCAAG	2779
QY	901	ACCGTGGCGCGGAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	960
Db	2780	ACCGTGGCGCGGAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	2839
QY	961	GTGCAGAACGCAACCCGAGTGCAGACCATCTGCGCGCTCTCGGCGCCCGCGCCAGC	1020
Db	2840	GTGCAGAACGCAACCCGAGTGCAGACCATCTCTGCGCGCTCTCGGCGCCCGCGCCAGC	2899
QY	1021	CTGGAGGAGATGATGACCGCTGCCAGGGGTGGCGGCGCCGAGCCACAGGCGCGGTG	1080
Db	2900	CTGGAGGAGATGATGACCGCTGCCAGGGGTGGCGGCGCCGAGCCACAGGCGCGGTG	2959

QY	1081	CTGGCGCGCGGATGAGCGAGG--CCAACACACAGCGGTGATGATGCAAGAGCAACTTC	1137
Db	2960	CTGGCGCGCGGATGAGCGAGGTTGAACAACACCAACATCATGATGCAAGCGGAGCACTGC	3039
QY	1138	AAGGGCCCGCGGCGCATCGTCAAGTGTCTCAACTGCGGCAAGGAGGCGCACATCGCCGC	1197
Db	3020	AAGGGCCCGCAAGCGCACCATCAAGTGTCTCAACTGCGGCAAGGAGGCGCACCTGGCCGC	3079
QY	1198	AACGCGCGCGCCCGCGCAAGAGGGCTGCTGGAAGTGGCGCAAGGAGGCGCACAGATG	1257
Db	3080	AACGCGCGCGCCCGCGCAAGAGGGCTGCTGGAAGTGGCGCAAGGAGGCGCACAGATG	3139
QY	1258	AAGGACTGCACCGAGCGCGCAGGCGCAACTTCTGGAAGTGGCGCAAGGAGGCGCACAGGCG	1317
Db	3140	AAGGACTGCACCGAGCGCGCAGGCGCAACTTCTGGAAGTGGCGCAAGGAGGCGCACAGGCG	3194
QY	1318	CGCCCGCGCAACTTCTGCGAGCGCGCGCGGAGGCGCGGAGGAGGAGGCGCGCGGAGGCTTC	1377
Db	3195	AGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGCTTC	3254
QY	1378	CGCTTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1425
Db	3255	AGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	3314
QY	1426	GAGACCTGACCGCGCTGAGAGGCGCTGTCGCGCAACGAGGAGGAGGAGGAGGAGGAGG	1479
Db	3315	TATCCTTTAGCTTCCCTCAGATCAGTCTTTGCGAGGAGGAGGAGGAGGAGGAGGAGG	3368
RESULT 13			
AX149648			
LOCUS	AX149648	4288 bp	DNA linear PAT 08-JUN-2001
DEFINITION	Sequence 2 from Patent WO0136614.		
ACCESSION	AX149648		
VERSION	AX149648.1 GI:14348047		
KEYWORDS	Human immunodeficiency virus		
SOURCE	Human immunodeficiency virus		
ORGANISM	Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate lentivirus group.		
REFERENCE	1 Shao, Y., Wagner, R., Wolf, H. and Graf, M.		
AUTHORS	The genome of the hiv-1 inter-subtype (c/h') and use thereof		
TITLE	Patent: WO 0136614-A 2 25-MAY-2001;		
JOURNAL	Geneart GMBH Gesellschaft fuer angewandte Biotechnologie (DE);		
FEATURES	Shao, Yiming (CN)		
source	Location/Qualifiers		
	1..4288		
	/organism="Human immunodeficiency virus"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:12721"		
BASE COUNT	1026 a	1358 c	1397 g
ORIGIN	507 t		
Query Match 81.5%; Score 1205.4; DB 6; Length 4288;			
Best Local Similarity 89.5%; Pred. No. 8.5e-122;			
Matches 1335; Conservative 0; Mismatches 141; Indels 15; Gaps 3;			
QY	1	ATGGGCGCGCGCGCCAGCATCTCTGCGGCGGCGCAAGCTGGAGCGCTGGAGCGCATCCGC	60
Db	13	ATGGGCGCGCGCGCCAGCATCTCTGAGGGCGGCAAGCTGGACAAGTGGAGAAGATCAGG	72
QY	61	CTGGCGCCCGCGGCGCAAGAGTGTCTACATGATGAAGCACTGGTGTGGCGCAGCGGAG	120
Db	73	CTGAGCGCCCGCGGCAAGAACCACTACATGCTGAAGCACTGGTGTGGCGCAGCGGAG	132
QY	121	CTGGAGAGTTGCGCCCTGAACCCCGGCTGCTGGAGACCGAGGAGGCTGCAAGCAGATC	180
Db	133	CTGGAGAGTTGCGCCCTGAACCCCGGCTGCTGGAGACCGAGGAGGCTGCAAGCAGATC	192
QY	181	ATCGCGCAGTGCACCCCGCTGCAGACCGGAGCGGAGGAGCTGAAGAGCCTGTTCAAC	240
Db			

```
Db 193 ATGAAGCAGCTGCAGAGCGCCCTGCAGACCGGCACCCAGGAGCTGAGGAGCCCTGTTCAAC 252
QY 241 ACCGTGGCACCCCTGTACGCGGTGCAGAGAGATGAGGTCGGGACACACCAAGAGGCG 300
Db 253 ACCGTGGCACCCCTGTACTCGTGCACACCGAGATGAGGTCAGGTCAGGTCAGGTCAGG 312
QY 301 CTGGACAAGATCGAGGAGGAGCAGAACAAAGTGCAGCAGAGAAGATCCAGCAGGCGCAGGCG 360
Db 313 CTGGACAAGATCGAGGAGGAGCAGAACAAAGATCCAGCAGAGAAGATCCAGCAGGCGCAGGAG 372
QY 361 GCCGACAAAGGCAAGGTGAGCCAGAACTACCCCATCTGTCAGAAACCTGCGAGGCGCAGATG 420
Db 373 GCCGAC---GGCAAGGTGAGCCAGAACTACCCCATCTGTCAGAAACCTGCGAGGCGCAGATG 429
QY 421 GTGCACACAGCCATCAGCCCGCCGACCCCTGAACGCCCTGGGTGAAGTGTATCGAGGAGAAG 480
Db 430 GTGCACACAGCCATCAGCCCGCAGGACCCCTGAATGTCATGGGTGAAGTGTGAGGAGAAG 489
QY 481 GCCITCAGCCCGAGGTGATCCCATGTTCAACGCCCTGAGCGAGGCGCCACCCCGCAG 540
Db 490 GCCTTCAGCCCGAGGTGATCCCATGTTCAACGCCCTGAGCGAGGCGCCACCCCGCAG 549
QY 541 GACCTGAACACGATGTTGAACACCGTGGCGGCCACAGCGCGCCATGTCAGATGCTCAAG 600
Db 550 GACCTGAACACGATGCTGAACACCGTGGCGGCCACAGCGCGCCATGTCAGATGCTCAAG 609
QY 601 GACACCATCAACAGGAGGCGCGCGAGTGGGACCGGTGCACCCCGTGCACCGCGCGCC 660
Db 610 GACACCATCAACAGGAGGCGCGCGAGTGGGACAGGCTGCACCCCGTGCACCGCGCGCC 669
QY 661 ATCGCCCCCGCCAGATGCGCGGAGCCCGCGGCGAGCAGATCGCGGCGCCACACAGCAGC 720
Db 670 ATCGCCCCCGCCAGATGAGGGAGCCCGCGGCGAGCAGATCGCGGCGCCACACAGCAGC 729
QY 721 CTGCAGGAGCAGATCGCCTGGATGACCAACACCCCGCATCCCGTGGGGCGACATCTAC 780
Db 730 CTGCAGGAGCAGATCGCCTGGATGACCAACACCCCGCGTGCCTGGGGCGACATCTAC 789
QY 781 AAGCGGTGATCATTCCIGGCGCCIGAACACAGATCGTGGCGGATGTACAGCCCGGAGCAGTC 840
Db 790 AAGAGGTGATCATCTCTGGTGTAAACACAGATCGTGGAGATGTACAGCCCGCAGCAGTC 849
QY 841 CTGGACATCAAGCAGGCGCGCCCAAGGAGCGCTTCGCGGACTACGTGAGCCCGCTTCTCAAG 900
Db 850 CTGGACATCAAGCAGGCGCGCCCAAGGAGCGCTTCAGGAGTACGTGAGCAGGTTCTCAAG 909
QY 901 ACCCTGGCGCGGAGCAGGAGCAGCAGCAGGAGTGAAGACTGGATGATGACCGACACCCCTG 960
Db 910 ACCCTGAGGCGGAGCAGGAGCAGCAGCAGGCGGTGAAGACTGGATGACCGACACCCCTG 969
QY 961 GTGCAGAACGCCAACCCGACTGCAAGACCACTCCTGCGCGCTCTCGCGCCCGCGCGCAGC 1020
Db 970 GTGCAGAACGCCAACCCGACTGCAAGACCACTCCTGAGGCGCTCTGCGCGCCCGCGCGCAGC 1029
QY 1021 CTGGAGGAGATGATGACCGCTGCGCAGGCGGTGGCGCGCCCGCAGCAGCAGGCGCGCTG 1080
Db 1030 ATCGAGGAGATGATGACCGCTGCGCAGGCGGTGGCGCGCCCGCAGCAGCAGGCGCGCTG 1089
QY 1081 CTGGCCGAGGAGATGAGCCAGGCCAACACACGCGTGTATGATGACAGAGAGCAACTCAAG 1140
Db 1090 CTGGCCGAGGAGATGAGCCAGGCCAACACACGCGCTGTATGATGACAGAGAGCAACTCAAG 1149
QY 1141 GCCCGCCCGGCGCATGTTAAGTGTCTCAAGTGTGGGAGGAGGCGCAGATGCGCGCGCAAG 1200
Db 1150 GGCAGCAAGAGATCGTGAAGTGTCTCAAGTGTGGGAGGAGGCGCAGATGCGCGCGCAAG 1209
QY 1201 TGCCCGCGCCCGCAGAGAGGCTGCTGGAAGTGCAGGAGGAGGCGCAGAGTGAAG 1260
Db 1210 TGCAAGGCGCCCGCAGAGAGGCTGCTGGAAGTGCAGGAGGAGGCGCAGAGTGAAG 1269
QY 1261 GACTGCACCGAGCGCCAGGCGCAACTTCTGGGCAAGATCTGGCCCGCAGCAGGCGCGC 1320
Db 1270 GACTGCACCGAGCGCAGGCGCAACTTCTGGGCAAGATCTGGCCCGCAGCAGGCGCGC 1329
```

```
QY 1321 CCGGGCAACTTCTCTGACAGCGCGCGCGGAGCCACCGCCCGCGCGGAGAGCTTCCGC 1380
Db 1330 CCGGGCAACTTCTCTGACAAACAGCGCGCGGAGCCACCGCCCGCGCGGAGAGCTTCCGC 1389
QY 1381 TTCGAGGAG-----ACCACCCCGCGCGCGCAGAGCAGGAG-----ASCAAGGAGCGCGGAG 1428
Db 1390 TTCGAGGAGGAGACCAACCCCGCGCGCAGAGCAGGAGCCCATCGACAAGGAGCTGTAC 1449
QY 1429 ACCCTGACGAGCCTGAGAGAGCCTGTGCGGCAACGACCCCGCTGAGCCAGTAA 1479
Db 1450 CCCTGACGAGCCTGAGAGAGCCTGTGCGGCAACGACCCCGCAGCAGCAGGAA 1500
```

RESULT 14
AF201927

LOCUS AF201927 1509 bp DNA linear SYN 16-MAR-2000
DEFINITION Synthetic construct gag protein gene, complete cds.

ACCESSION AF201927

VERSION AF201927.1 GI:7248702

KEYWORDS

SOURCE synthetic construct

ORGANISM

synthetic construct

artificial sequences.

REFERENCE 1 (bases 1 to 1509)

AUTHORS

zur Megede,J., Chen,M.C., Doe,B., Schaefer,M., Greer,C.E.,

Seiby,M., Otten,G.R. and Barnett,S.W.

Increased expression and immunogenicity of sequence-modified human

immunodeficiency virus type 1 gag gene

J. Virol. 74 (6), 2628-2635 (2000)

MEDLINE 20148954

PUBMED 10684277

REFERENCE 2 (bases 1 to 1509)

zur Megede,J. and Barnett,S.W.

AUTHORS

Direct Submission

Submitted (04-NOV-1999) Vaccines, Chiron Corporation, 4560 Horton,

Emeryville, CA 94608, USA

JOURNAL

Location/Qualifiers

FEATURES

source

CDS

```
1..1509
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
1..1509
/note="HIV-1SF2 p55gag"
/codon_start=1
/transl_table=11
/product="gag protein"
/protein_id="AAF43628.1"
/db_xref="GI:7248703"
translation="MGARASVLSGGELDKWEKIRLRPGKKKKYKLRHIVWASRELEHF
AVNPLLETSGCRQILQLPQLSLOIGSEELRSLYNTVATLYCVHQRIDVDKIDKEALE
KIEEQNKSKKKAQAAAGTGNSSQVSNYPVQLQGMVHQAIPTLNARVYK
VEKAFSPVPMFSALSEGATPODLNLTMLNTVGGHOAOMLKETINEEAAEWDRVH
PVHAGPIAPGMREPRGSDIAGTTSTLQEOIGWMTNPPVPEIYKRWIILGNLIV
RMVSPSTSLDIRQPKPFYDFYKTLRAEQASODYKNWMIETLIVQANPDKTV
ILKALGPAATLEEMTACGCGGPGHKARVLAEMSOTNPATIMMORGNFRNQRKV
KCFNCGKEGHTARNCRAPKKGCRGREGHOMKDCITPCANFLGKIWPSYKGRPGNF
LQSRPEPTAPPEESFRFGHEKTPSQKQEPIDKELYPLTSLRSLFGNDPSSQ"
```

BASE COUNT 328 a 543 c 479 g 159 t

ORIGIN

Query Match 81.0%; Score 1197.8; DB 12; length 1509;
Best local similarity 89.3%; Pred. No. 7.6e-121;
Matches 1347; Conservative 0; Mismatches 132; Indels 30; Gaps 4;

```
QY 1 ATGGGGCGCGCGCGCAGCATCCTGCGCGCGCGGCAAGCTGGAGCGCTGGGAGCGCATCCGC 60
Db 1 ATGGGGCGCGCGCGCAGCATCCTGCTGAGCGCGCGCGGAGCTGGACAATGGGAGAGATCCGC 60
QY 61 CTGGCGCGCGCGCGGCGAAGAAGTGTACATGATGAGGACACCTGGTGTGGCGCAGCGCGGAG 120
Db 61 CTGGCGCGCGCGCGGCGAAGAAGTGTACATGATGAGGACACATCGTGTGGCGCAGCGCGGAG 120
```


QY 121 CTGGAGAGTTCGCTGAACCCCGGCTGCTGGAGACCAAGAGGGCTGCAAGCAATC 180
Db 121 CTGGAGGCTTGGCGTGAACCCCGGCTGCTGGAGACCAAGAGGGCTGCGGCGAGATC 180
QY 181 ATCCGCCAGTGCACCCCGCTGCAGACCGGCGAGGAGCTGAGAGGCTGTTCAC 240
Db 181 CTGGGCCAGTGCAGCCCGCTGCAGACCGGCGAGGAGCTGCGGCGAGCTGTACAC 240
QY 241 ACCCTGGCCACCTGTACTGCTGCAGAGAGAGATCGAGTCCGGACACCAAGAGGCG 300
Db 241 ACCGTGGCCACCTGTACTGCTGCAGACCGGCGATCGAGCTCAAGGACACCAAGAGGCG 300
QY 301 CTGGACAAGATCGAGGAGGAGCAGACCAAGTCCAGCAGAGATCCAGCAGGCGCGAGGCG 360
Db 301 CTGGAGAAGATCGAGGAGGAGCAGACCAAGTCCAGCAGAGGCGCGAGGCGCGCGGCG 360
QY 361 GCCG-----ACAGGGCAAGTGTAGCCAGAACTACCCCATCTGTCAGAACCTG 408
Db 361 GCCCGCGGACCGGCAACACAGCCAGGTGAGCCAGAACTACCCCATCTGTCAGAACCTG 420
QY 409 CAGGCCAGATGGTGACCAAGGCCATCAGCCCCCGCACCTCTGAACGGCTGGTGAAGGTG 468
Db 421 CAGGCCAGATGGTGACCAAGGCCATCAGCCCCCGCACCTCTGAACGGCTGGTGAAGGTG 480
QY 469 ATCGAGGAGAGGCTTCAGCCCCCGAGGTGATCCCCATGTTACCGCGCTGAGCGAGGCG 528
Db 481 GTGGAGGAGAGGCTTCAGCCCCCGAGGTGATCCCCATGTTACCGCGCTGAGCGAGGCG 540
QY 529 GCCACCCCGCAGGACCTGACACAGATGTTGAACACCGTGGCGGCGCCAGCGCCCATG 588
Db 541 GCCACCCCGCAGGACCTGACACAGATGTTGAACACCGTGGCGGCGCCAGCGCCCATG 600
QY 589 CAGATGCTGAAGGACACCATCAACGAGGAGGCGCGCGGAGTGGGACCGCTGCACCGGCG 648
Db 601 CAGATGCTGAAGGAGACCATCAACGAGGAGGCGCGCGGAGTGGGACCGCTGCACCGGCG 660
QY 649 CACGCGGCGCCATCGCCCGCGGCGAGATGCGCGAGCGCGCGGAGGACATCGCGGCG 708
Db 661 CACGCGGCGCCATCGCCCGCGGCGAGATGCGCGAGCGCGCGGAGGACATCGCGGCG 720
QY 709 ACCACGAGCCCTGCAGGAGCAGATCGGCTGGATGACAGCAACCCCGCCATCGCGGCG 768
Db 721 ACCACGAGCCCTGCAGGAGCAGATCGGCTGGATGACCAACAACCCCGCCATCGCGGCG 780
QY 769 GCGGACATCTACAAGCGGTGGATCATCTGGGCGCTGAACAAGATCGTGGGATGACAG 828
Db 781 GCGGAGATCTACAAGCGGTGGATCATCTGGGCGCTGAACAAGATCGTGGGATGACAG 840
QY 829 CCGGTGAGCATCTGACATCAAGCGGCGGCGGAGGCGGCGGAGGCGCTCCGCGACTACG 888
Db 841 CCGACGATCTCTGACATCGGCGGCGGAGGCGGCGGAGGCGCTCCGCGACTACG 900
QY 889 CGCTTCTTAAGACCCCTGCGGCGGAGCAGACCCAGGAGGTGAAGAAGTGGATGAC 948
Db 901 CGCTTCTTAAGACCCCTGCGGCGTGAAGCGGCGGAGGAGGTGAAGAAGTGGATGAC 960
QY 949 GACACCCCTGCTGGTGAGAGGCGGCGGAGGCGGAGGCGGAGGCGGCGGCGGCGGCG 1008
Db 961 GAGACCCCTGCTGGTGAGAGGCGGCGGAGGCGGAGGCGGAGGCGGCGGCGGCGGCG 1020
QY 1009 CCGGCGGCGGAGGAGGAGTGTAGCGGCGCTGCCAGGCGGCGGCGGCGGCGGCGGCGG 1068
Db 1021 CCGGCGGCGGAGGAGGAGTGTAGCGGCGCTGCCAGGCGGCGGCGGCGGCGGCGGCGG 1080
QY 1069 AAGGCGGCGGCTGCGGCGGAGGCGGATGAGCGAGGCGGCGGCGGCGGCGGCGGCGG 1122
Db 1081 AAGGCGGCGGCTGCGGCGGAGGCGGATGAGCGAGGCGGCGGCGGCGGCGGCGGCGG 1140
QY 1123 CAGAAGAGCAACTTCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1182
Db 1141 CAGCGCGGCAACTTCCGCAACAGCGGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200
QY 1183 GGCCACATCGCGCGCAACTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1242

Db 1201 GGCCACACCGCCAGGAAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1260
QY 1243 GAGGGCCACAGATGAAGGACTGCACGAGCGCCAGGCGCAACTTCTCGGCAAGATCTGG 1302
Db 1261 GAGGGCCACAGATGAAGGACTGCACGAGCGCCAGGCGCAACTTCTCGGCAAGATCTGG 1320
QY 1303 CCCAGCCACAAGGGCGCGCGGCGCAACTTCTCGAGAGCGCCGAGCGCCACCGCGCGCC 1362
Db 1321 CCCAGCTACAAGGGCGCGCGGCGCAACTTCTCGAGAGCGCCGAGCGCCACCGCGCGCC 1380
QY 1363 CCGCGCGAGAGCTTCGCTTC-----GAGGAGACCGCCCGCGGCGCAAGCAGGAG--- 1413
Db 1381 CCGGAGGAGAGCTTCGCTTCGCGGAGGAGAGACCGCCCGGAGGCGGCGGCGGCGGCGG 1440
QY 1414 ---AGCAAGGAGCGGAGACCTGACAGCTGAAGAGAGCTGTCGCAACGACCCCGCTG 1470
Db 1441 ATCGACAAGGAGCTGTACCCCTGACCGAGCTGCGCAGCTGTCGCAACGACCCCGAGC 1500
QY 1471 AGCCAGTAA 1479
Db 1501 AGCCAGTAA 1509
RESULT 15
AF287354
LOCUS Synthetic construct HIV-1-derived gag protein gene, complete cds.
DEFINITION AF287354
ACCESSION AF287354
VERSION AF287354.1 GI:11066866
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 1548)
AUTHORS Fuller, M. and Anson, D.S.
TITLE Helper plasmids for production of HIV derived vectors
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1548)
AUTHORS Fuller, M. and Anson, D.S.
TITLE Direct Submission
JOURNAL Submitted (12-JUL-2000) Chemical Pathology, Women's and Children's
Hospital, 72 King William Road, North Adelaide, SA 5006, Australia
FEATURES
Location/Qualifiers
1..1548
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="derived from Human immunodeficiency virus type 1
isolate YU-2"
37..1539
/note="codon optimized reading frame"
/codon_start=1
/transl_table=11
/product="gag protein"
/protein_id="AAG28738.1"
/db_xref="GI:11066867"
/translation="MGARASVLSAGELDKWEIKRLPGGKKQYRLKHHVWASRELERF
AVDPGILETSEGQILGOLQPSLQTSFEELSLYNTVATLYCVHQKIEVKDTKEALE
KIEEQNKSKKAAQAAADTGNSSQVSNYPVQNLOGQVHQAI SPRTLNAAWKVVE
EKAFSPVPMFSALSEGATPDQNTMLNTVGHOAAQMLKETINEEAEDWRLHPV
HAGPIAQOMREPRGSDIAGTTSILQEQIGWNTNPPIPVGEIYKRWIILGLNKIVRM
YSPTSLDIROGPKPEFRDYVDYFYKTLRAEQASOEYKNNWMTETLLVQNPANPDCKTIL
KALPAATLEEMTACQGVGGGCHKARVLAEMSQVTSNATIMMQRFNQRKTVKC
FNCGKEGHIANKRAPRKKGCKGKQEGHMKDCTERQANFTGKIWPSHKRGRPGNFLO
SRPEPTAPSEESVRFGEETTTTSQKQEPIDKELYPLASLRSFGSDPSSQ"

BASE COUNT 374 a 495 c 515 g 164 t
ORIGIN
Query Match 78.4%; Score 1159.6; DB 12; Length 1548;
Best Local Similarity 87.8%; Pred. No. 1e-116;
Matches 1319; Conservative 0; Mismatches 159; Indels 24; Gaps 4;

Qy	1	ATGGGGCGCCGCGCAGCATCTCTGGCGCGCGGCAAGCTGGACGCCCTGGAGCGCATCGC	60
Db	37	ATGGGCGCAGGGCCAGCGCTGCTGAGCGCGCGGAGTGGACAAAGTGGAGNAGATCAGG	96
Qy	61	CTGCGCCGCGCGGCAAGAACTGCTACATGATGAAGCACTGGTGTGGGCCAGCGCGGAG	120
Db	97	CTGAGGCGCGCGGCAAGAAAGCAGTACAGGCTGAAGCACATCGTGTGGGCCAGCAGGAG	156
Qy	121	CTGGAGAACTTCGCCCTGAAACCCCGGCTGCTGGAGACAGCGAGGGCTGCAAGCAATC	186
Db	157	CTGGAGAGTTTCGCCCTGGACCCCGGCTGCTGGAGACAGCGAGGGCTGAGGAGCTGTACAA	216
Qy	181	ATCCGCCAGCTGCACCCCGGCCCTGCAGACCGGCGAGCGAGAGCTGAAGAGCCCTGTTCA	240
Db	217	CTGGGCCAGCTGCAGCCGAGCCCTGCAGACCGGCGAGCGAGGCTGAGGAGCTGTACAA	276
Qy	241	ACCGTGGCAACCCCTGTACTGGCTGCACGAGAGATCGAGGTCCGCGACACCAAGAGGCG	300
Db	277	ACCGTGGCACCCCTGTACTGGCTGCACCGAGAGATCGAGGTGAAGACACCAAGAGGCG	336
Qy	301	CTGGACAAGATCGAGGAGGAGCAGAACTACTGCCAGCAGAAATCCAGCATCCAGCAG--	354
Db	337	CTGGAGAAGATCGAGGAGGAGCAGAACACAGAGCAAGAGAAGAGGCCAGCAGGCCGCG	396
Qy	355	GAGCGCGCGCACAAAGGCGAAGGTGAGCCAGAACTACCCCATCGTGCAGAACTGCAGGCG	414
Db	397	GACACGGCAACAGCAGCGCAGGTGAGCCAGAACTACCCCATCGTGCAGAACTGCAGGCG	456
Qy	415	CAGATGTCACACGAGGCCATCAGCCCCCGCACCCCTGAACGCCCTGGTGAAGGTGATGAG	474
Db	457	CAGATGTCACACGAGGCCATCAGCCCCCGCACCCCTGAACGCCCTGGTGAAGGTGATGAG	516
Qy	475	GAGAAGGCCCTCAGCCCGAGGTGATCCCATGTTCAACGCCCTTGAGCAGGGCGGCGAC	534
Db	517	GAGAAGGCCCTCAGCCCGAGGTGATCCCATGTTCAACGCCCTTGAGCAGGGCGGCGAC	576
Qy	535	CCCCAGGACCTGAACACGATGTTGAACACCGTGGCGGCGCACCCAGCCGATGCAGATG	594
Db	577	CCCCAGGACCTGAACACGATGTTGAACACCGTGGCGGCGCACCCAGCCGATGCAGATG	636
Qy	595	CTGAAGGACACCATCAACGAGGAGGCGCGGAGTGGGACCGCGTGCACCCCGTGCACGCC	654
Db	637	CTGAAGGAGACCATCAACGAGGAGGCGCGGAGTGGGACCGCGTGCACCCCGTGCACGCC	696
Qy	655	GGCCCCATCGCCCCCGGCGACATGGCGAGCGCCCGGCGACGACATCGCGGCGCACCCAC	714
Db	697	GGCCCCATCGCCCCCGGCGACATGGCGAGCGCCCGGCGACGACATCGCGGCGCACCCAC	756
Qy	715	AGCACCTGCAGGAGCAGATCGCCTGGATGACCAAGCAACCCCCCATCCCGTGGCGAC	774
Db	757	AGCACCTGCAGGAGCAGATCGCCTGGATGACCAAGCAACCCCCCATCCCGTGGCGAC	816
Qy	775	ATCTACAAGCGGTGGATCATCTCGGCGCTGAACAAGATCGTGGGATGTACAGCCCCGTC	834
Db	817	ATCTACAAGAGGTGGATCATCTCGGCGCTGAACAAGATCGTGGGATGTACAGCCCCAC	876
Qy	835	AGCATCTGGACATCAAGCAGGCGCCCAAGAGCCCTTCGCGGACTACGTGCACCGCTTC	894
Db	877	AGCATCTGGACATCAAGCAGGCGCCCAAGAGCCCTTCGCGGACTACGTGCACCGCTTC	936
Qy	895	TCAAGACCCCTCGCGCGGCGAGCAGACCCAGGAGGTGAAGAACTGATGACCGACACG	954
Db	937	TACAAGACCCCTGAGGCGCGGAGCAGGCCAGGAGGTGAAGAACTGGATGATCCAGAC	996
Qy	955	CTGCTGTTGCAGAACGCCAACCCCGACTGCAAGACCATCTTCGCGGCTCTCGGCCCGCG	1014
Db	997	CTGCTGTTGCAGAACGCCAACCCCGACTGCAAGACCATCTTCGCGGCTCTCGGCCCGCG	1056
Qy	1015	GCCAGCCTGGAGAGATGATGACCCGCTGCGAGGGCGGTGGCGGCGCCAGCCACAGGCG	1074
Db	1057	GCCAGCCTGGAGAGATGATGACCCGCTGCGAGGGCGGTGGCGGCGCCAGCCACAGGCG	1116
Qy	1075	CGCGTGTGCGCGAGCGGATGAGCCAGGCCAACACCCAGCG-----TGATGATGAGAG	1128

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2003, 14:10:43 : Search time 144.5 Seconds
(without alignments)
1021.419 Million cell updates/sec

Title: US-09-475-704A-2
Perfect score: 60
Sequence: 1 qacatcgccagggcccaaa.....tggacccttcttcagagacc 60

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1560708 seqs. 1229959015 residues

Total number of hits satisfying chosen parameters: 1177748

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_NA:

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
- 6: /cgn2_6/ptodata/2/pubpna/PCT_US_PUBCOMB.seq:
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
- 15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
- 16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
- 17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	21	35.0	87	9	US-09-864-761-28254
c 2	20.4	34.0	60	9	US-09-007-093-7
c 3	20.4	34.0	65	12	US-09-908-975-25099
c 4	20.2	33.7	65	12	US-09-908-975-30489
c 5	19.8	33.0	77	11	US-09-805-353-16
c 6	19.4	32.3	100	16	US-09-728-445-336
c 7	18.6	31.0	60	12	US-09-908-975-9744
c 8	18.4	30.7	60	12	US-09-908-975-13339
c 9	18.2	30.3	60	12	US-09-908-975-7983
c 10	18.2	30.3	88	12	US-10-133-973-102
c 11	18.2	30.3	89	10	US-09-971-980-18
c 12	18	30.0	65	12	US-09-908-975-30263
c 13	17.8	29.7	56	9	US-09-967-013-81
c 14	17.8	29.7	56	9	US-09-967-013-82
c 15	17.8	29.7	56	11	US-09-863-733A-80
c 16	17.8	29.7	56	11	US-09-863-733A-81

c 17	17.8	29.7	56	14	US-10-116-420-80	Sequence 80, Appl
c 18	17.8	29.7	56	14	US-10-116-420-81	Sequence 81, Appl
c 19	17.8	29.7	59	9	US-09-967-013-88	Sequence 88, Appl
c 20	17.8	29.7	59	9	US-09-967-013-89	Sequence 89, Appl
c 21	17.8	29.7	59	11	US-09-863-733A-87	Sequence 87, Appl
c 22	17.8	29.7	59	11	US-09-863-733A-88	Sequence 88, Appl
c 23	17.8	29.7	59	11	US-09-863-733A-93	Sequence 93, Appl
c 24	17.8	29.7	59	11	US-09-863-733A-94	Sequence 94, Appl
c 25	17.8	29.7	59	14	US-10-116-420-87	Sequence 87, Appl
c 26	17.8	29.7	59	14	US-10-116-420-88	Sequence 88, Appl
c 27	17.8	29.7	59	14	US-10-116-420-93	Sequence 93, Appl
c 28	17.8	29.7	59	14	US-10-116-420-94	Sequence 94, Appl
c 29	17.8	29.7	60	12	US-09-908-975-13021	Sequence 13021, A
c 30	17.8	29.7	60	12	US-09-908-975-13510	Sequence 13510, A
c 31	17.8	29.7	60	12	US-09-908-975-16956	Sequence 16956, A
c 32	17.8	29.7	60	12	US-09-908-975-20072	Sequence 20072, A
c 33	17.8	29.7	63	9	US-09-811-672-4	Sequence 4, Appli
c 34	17.8	29.7	63	12	US-10-161-660-4	Sequence 4, Appli
c 35	17.8	29.7	65	12	US-09-908-975-1045	Sequence 1045, Ap
c 36	17.8	29.7	65	12	US-09-908-975-24242	Sequence 24242, A
c 37	17.8	29.7	90	10	US-09-971-798-6	Sequence 6, Appli
c 38	17.6	29.3	50	12	US-10-322-360-119	Sequence 119, App
c 39	17.6	29.3	56	9	US-09-760-574-10	Sequence 10, Appl
c 40	17.6	29.3	56	11	US-09-766-442A-10	Sequence 8727, Ap
c 41	17.6	29.3	60	12	US-09-908-975-8727	Sequence 1201, Ap
c 42	17.6	29.3	85	9	US-09-815-242-1201	Sequence 21, Appl
c 43	17.4	29.0	30	10	US-09-777-430A-21	Sequence 296, App
c 44	17.4	29.0	30	11	US-09-864-636A-296	Sequence 164, App
c 45	17.4	29.3	30	12	US-09-758-282-164	

ALIGNMENTS

RESULT 1

- US-09-864-761-28254/c
- Sequence 28254, Application US/09864761
- Patent No. US20020048763A1
- GENERAL INFORMATION:
- APPLICANT: Penn, Sharon G.
- APPLICANT: Rank, David R.
- APPLICANT: Hanzel, David K.
- APPLICANT: Chen, Wensheng
- TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL
- TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
- FILE REFERENCE: Aomic-X-1
- CURRENT APPLICATION NUMBER: US/09/864,761
- CURRENT FILING DATE: 2001-05-23
- PRIOR APPLICATION NUMBER: US 60/180,312
- PRIOR FILING DATE: 2000-02-04
- PRIOR APPLICATION NUMBER: US 60/207,456
- PRIOR FILING DATE: 2000-05-26
- PRIOR APPLICATION NUMBER: US 09/632,366
- PRIOR FILING DATE: 2000-08-03
- PRIOR APPLICATION NUMBER: GB 24263.6
- PRIOR FILING DATE: 2000-10-04
- PRIOR APPLICATION NUMBER: US 60/236,359
- PRIOR FILING DATE: 2000-09-27
- PRIOR APPLICATION NUMBER: PCT/US01/00666
- PRIOR FILING DATE: 2001-01-30
- PRIOR APPLICATION NUMBER: PCT/US01/00667
- PRIOR FILING DATE: 2001-01-30
- PRIOR APPLICATION NUMBER: PCT/US01/00664
- PRIOR FILING DATE: 2001-01-30
- PRIOR APPLICATION NUMBER: PCT/US01/00669
- PRIOR FILING DATE: 2001-01-30
- PRIOR APPLICATION NUMBER: PCT/US01/00665
- PRIOR FILING DATE: 2001-01-30
- PRIOR APPLICATION NUMBER: PCT/US01/00668
- PRIOR FILING DATE: 2001-01-30
- PRIOR APPLICATION NUMBER: PCT/US01/00663
- PRIOR FILING DATE: 2001-01-30
- PRIOR APPLICATION NUMBER: PCT/US01/00662

```
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,406
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 28254
; LENGTH: 87
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004870.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
; OTHER INFORMATION: SWISSPROT HIT: P75080, EVALUE 5.60e-00
US-09-864-761-28254

Query Match 35.0%; Score 21; DB 9; Length 87;
Best Local Similarity 73.0%; Pred. No. 3.9e-02;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 20 AGGAGCCCTTCGGGACTACGTGGACCGCTTCTTCAAA 56
      ||||| || |||| || |||| || |||| || ||||
Db 86 AGGAGCCCTTCGAACACTACCTGAACCATGATCAA 50

RESULT 2
US-09-007-093-7
; Sequence 7, Application US/09007093
; Patent No. US2002025315A1
; GENERAL INFORMATION:
; APPLICANT: Anand, Naveen N
; APPLICANT: Barber, Brian H
; APPLICANT: Cates, George A
; APPLICANT: Caterina, Judith E
; APPLICANT: Klein, Michel H
; TITLE OF INVENTION: CHIMERIC ANTIBODIES FOR DETECTION OF
; TITLE OF INVENTION: ANTIGENS TO SELECTED CELLS OF THE IMMUNE SYSTEM
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/007,093
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/483,576
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-765

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 60 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-007-093-7

Query Match 34.0%; Score 20.4; DB 9; Length 60;
Best Local Similarity 65.2%; Pred. No. 6.3e+02;
Matches 30; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 13 GGCCCAAGGAGCCCTTCGGGACTACGTGGACCGCTTCTTCAAGA 58
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 13 GGTCTAAAGAACCTTTTAGAGACTATGTGTAGAGTTTATATAAGA 58

RESULT 3
US-09-908-975-25099
; Sequence 25099, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPL
; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 25099
; LENGTH: 65
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-908-975-25099

Query Match 34.0%; Score 20.4; DB 12; Length 65;
Best Local Similarity 61.1%; Pred. No. 6.3e+02;
Matches 33; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 6 CGGCCAGGGCCCAAGGAGCCCTTCGGGACTACGTGGACCGCTTCTTCAAGAC 59
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2 CCCCAGACACCCCAAGGAGATGCTTACCACCGCTGCACCGACATGTC 55

RESULT 4
US-09-908-975-30489
; Sequence 30489, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPL
; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
```


Query Match 30.7%; Score 18.4; DB 12; Length 60;
Best Local Similarity 59.6%; Pred. No. 3.1e+03;
Matches 31; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 3 CATCCGCGAGGCCCCCAAGGAGCCCTTCGCGGACTAGCTGGACCGCTTC 54
DB 6 CATCAGCAACCTCTTCAGAGGAGGACTGCCACCAGAGAGATCGATGACCTTC 57

RESULT 9
US-09-908-975-7983
; Sequence 7983, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7983
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-908-975-7983

Query Match 30.3%; Score 18.2; DB 12; Length 60;
Best Local Similarity 58.2%; Pred. No. 3.6e+03;
Matches 32; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 5 TCCGCGAGGCCCCCAAGGAGCCCTTCGCGGACTAGCTGGACCGCTTC 59
DB 6 TCAGCTGGTCTCAGCAAGGATGCCCGCCGCAAGAGGAGCGCGGAGCTCTCCACAC 60

RESULT 10
US-10-133-973-102/c
; Sequence 102, Application US/10133973
; Publication No. US20030149254A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David
; TITLE OF INVENTION: METHODS AND COMPOSITIONS COMPRISING RENILLA GFP
; FILE REFERENCE: A-68531-3/RMS/CYO
; CURRENT APPLICATION NUMBER: US/10/133,973
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 60/290,287
; PRIOR FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: US 09/710,058
; PRIOR FILING DATE: 2000-11-10
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 102
; LENGTH: 88
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: primer
US-10-133-973-102

Query Match 30.3%; Score 18.2; DB 12; Length 88;
Best Local Similarity 58.2%; Pred. No. 3.6e+03;
Matches 32; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 6 CCGCCAGGCGCCCAAGGAGCCCTTCGCGGACTAGCTGGACCGCTTC 60
DB 88 COTCCAGGTCACCTGTGAGCTCATCACCTCCTGCAGCGCGGTGTTCTTCAGGATC 34

RESULT 11
US-09-971-980-18/c
; Sequence 18, Application US/09971980
; Patent No. US20020164349A1
; GENERAL INFORMATION:
; APPLICANT: Weiner, David B.
; APPLICANT: Yang, Joo-Sung
; TITLE OF INVENTION: Compositions and Methods of Using Capsid Protein From Flaviviruses
; TITLE OF INVENTION: Pestiviruses
; FILE REFERENCE: Upn-4105
; CURRENT APPLICATION NUMBER: US/09/971,980
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/237,885
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 89
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: No. US20020164349A1e1 Sequence
US-09-971-980-18

Query Match 30.3%; Score 18.2; DB 10; Length 89;
Best Local Similarity 61.7%; Pred. No. 3.6e+03;
Matches 29; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 3 CATCCGCGAGGCCCCCAAGGAGCCCTTCGCGGACTAGCTGGACCGCT 49
DB 83 CTTCGCTTTCACCGCCATTGCCCCACCCGCGCGCTGCTGGACCGCT 37

RESULT 12
US-09-908-975-30263
; Sequence 30263, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPL
; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 30263
; LENGTH: 65
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-908-975-30263

Query Match 30.0%; Score 18; DB 12; Length 65;
Best Local Similarity 64.3%; Pred. No. 4.2e+03;
Matches 27; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 3 CATCCGCGAGGCCCCCAAGGAGCCCTTCGCGGACTAGCTGGACCGCT 44
DB 17 CATTTCACAGCCCTCCAGGAGAGCTTCAAGGGCTACGAAGA 58

Query Match:	29.7%	Score 17.8;	DB 11;	Length 56;
Best Local Similarity	62.2%	Pred. No. 5e-03;		
Matches	29;	Conservative	0;	Mismatches 17;
				Indels 0;
				Gaps 0;

Search completed: September 22, 2003, 17:30:58
Job time : 145.5 secs

RESULT 15
US-09-863-733A-80
; Sequence 80, Application: US/09863733A
; Publication No. US20030082537A1
; GENERAL INFORMATION:
; APPLICANT: Stantec, Jr., Vincent P.
; TITLE OF INVENTION: METHODS FOR GENETIC ANALYSIS OF DNA
; FILE REFERENCE: 11926-113001

```

RESULT 15
US-09-863-733A-80
; Sequence 80, Application: US/09863733A
; Publication No. US20030082537A1
; GENERAL INFORMATION:
; APPLICANT: Stantec, Jr., Vincent P.
; TITLE OF INVENTION: METHODS FOR GENERATING
; FILE REFERENCE: 11926-113001

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2003, 14:02:05 ; Search time 44 seconds
(without alignments)
501.886 Million cell updates/sec

Title: US-09-475-704A-2

Perfect score: 60

Sequence: 1 gacatccgcagggccccaa.....tgagccgctttctcaagacc 60

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 223691566 residues

Total number of hits satisfying chosen parameters: 230498

Minimum DB seq length: 0
Maximum DB seq length: 106

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA: *
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq: *
6: /cgn2_6/ptodata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	21.8	36.3	84	1	US-08-405-702A-7
2	20.8	34.7	84	1	US-08-405-702A-5
3	20.8	34.7	84	1	US-08-405-702A-9
4	20.4	34.0	60	3	US-08-943-136-7
5	20.4	34.0	60	4	US-08-973-518-7
6	20	33.3	72	1	US-08-303-275-192
7	20	33.3	72	1	US-08-303-275-193
8	19.6	32.7	57	6	5514566-17
9	18	30.0	38	3	US-09-282-147-28
10	18	30.0	72	2	US-08-707-237A-47
11	18	30.0	100	1	US-07-797-556-7
12	17.8	29.7	40	1	US-08-495-743-25
13	17.8	29.7	40	1	US-08-548-078-3
14	17.8	29.7	40	1	US-08-495-739-25
15	17.8	29.7	40	1	US-08-495-741-25
16	17.8	29.7	40	3	US-08-062-023-25
17	17.8	29.7	44	2	US-08-403-852D-35
18	17.8	29.7	44	3	US-08-510-646B-37
19	17.8	29.7	44	3	US-09-231-818-35
20	17.8	29.7	45	1	US-08-167-939A-14
21	17.8	29.7	45	1	US-08-567-538-14
22	17.8	29.7	63	3	US-08-750-419A-4
23	17.8	29.7	63	4	US-09-811-672-4
24	17.8	29.7	72	1	US-08-303-275-195
25	17.6	29.3	50	4	US-09-554-929-119
26	17.6	29.3	72	1	US-07-778-233B-72
27	17.6	29.3	72	1	US-07-963-321-72

28	17.6	29.3	72	1	US-08-290-641-72	Sequence 72, Appl
29	17.6	29.3	72	1	US-08-548-540-72	Sequence 72, Appl
30	17.6	29.3	72	5	PCT-US96-09809-72	Sequence 72, Appl
31	17.4	29.0	53	2	US-07-695-201B-13	Sequence 13, Appl
32	17.4	29.0	53	3	US-08-470-532-13	Sequence 13, Appl
33	17.2	28.7	48	1	US-08-151-574-14	Sequence 14, Appl
34	17.2	28.7	48	2	US-08-419-448-14	Sequence 14, Appl
35	17.2	28.7	48	4	US-09-233-510-14	Sequence 14, Appl
36	17.2	28.7	50	1	US-08-575-052-7	Sequence 7, Appl
37	17.2	28.7	50	1	US-08-614-516A-7	Sequence 7, Appl
38	17.2	28.7	50	1	US-08-770-557-7	Sequence 7, Appl
39	17.2	28.7	50	1	US-08-405-234-7	Sequence 7, Appl
40	17.2	28.7	65	3	US-09-191-852-16	Sequence 16, Appl
41	17.2	28.7	65	4	US-08-817-906-16	Sequence 16, Appl
42	17.2	28.7	65	5	PCT-US95-13376-16	Sequence 16, Appl
43	17.2	28.7	80	1	US-08-152-621-7	Sequence 7, Appl
44	17.2	28.7	80	1	US-08-306-691B-38	Sequence 38, Appl
45	17.2	28.7	80	5	PCT-US92-05035-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1
US-08-405-702A-7
; Sequence 7, Application US/08405702A
; Patent No. 5789389
; GENERAL INFORMATION:
; APPLICANT: Tarasewicz, Dariusz G
; APPLICANT: Schott, Brigitte
; APPLICANT: Holzmayer, Iatiana A.
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: BCL2 Derived Genetic Elements Associated
; TITLE OF INVENTION: with Sensitivity to Chemotherapeutic Drugs
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/405,702A
; FILING DATE: 17-MAR-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5789389nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 95,332
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; FIELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-405-702A-7

Query Match 36.3%; Score 21.6; DB 1; Length 84;
Best Local Similarity 61.4%; Pred. No. 1.3e+02;
Matches 35; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

3 CATCCGCCAGGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGAC 59

NAME: Stewart, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-733
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-943-136-7
Query Match 34.0%; Score 20.4; DB 3; Length 60;
Best Local Similarity 65.2%; Pred. No. 3.5e+02;
Matches 30; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
Qy 13 GGCCCAAGGAGCGCTTCGGCGACTACGTGGACCGCTTCTTCAAGA 58
Db 13 GGCTAAAGAACCTTTTAGAGACTATGTGATAGGTTTATATAAGA 58
RESULT 5
US-08-973-518-7
Sequence 7, Application US/08973518
Patent No. 6328962
GENERAL INFORMATION:
APPLICANT: Anand, Navreet N
APPLICANT: Barber, Brian H
APPLICANT: Gates, George A
APPLICANT: Caterini, Judith E
APPLICANT: Klein, Michel H
TITLE OF INVENTION: CHIMERIC ANTIBODIES FOR DELIVERY OF
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 6th Floor, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/973,518
FILING DATE: 07-APR-1998
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-781
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-973-518-7
Query Match 34.0%; Score 20.4; DB 4; Length 60;
Best Local Similarity 65.2%; Pred. No. 3.5e+02;
Matches 30; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
Qy 13 GGCCCAAGGAGCGCTTCGGCGACTACGTGGACCGCTTCTTCAAGA 58
Db 13 GGCTAAAGAACCTTTTAGAGACTATGTGATAGGTTTATATAAGA 58

Db 13 GGTCCTAAAGAACCTTTTAGAGACIATGTGTAGATAGGTTTATATAAGA 58
RESULT 6
US-08-303-275-192
Sequence 192, Application US/08303275
Patent No. 5766598
GENERAL INFORMATION:
APPLICANT: Paoletti, Enzo
APPLICANT: Tartaglia, James
APPLICANT: Cox, William I.
TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS RECOMBINANT
TITLE OF INVENTION: POXVIRUS VACCINE
NUMBER OF SEQUENCES: 205
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford
ADDRESSEE: c/o William S. Frommer
STREET: 530 Fifth Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/303,275
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/897,362
FILING DATE: 11-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2420
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 192:
SEQUENCE CHARACTERISTICS:
LENGTH: 72 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-303-275-192
Query Match 33.3%; Score 20; DB 1; Length 72;
Best Local Similarity 61.5%; Pred. No. 4.8e+02;
Matches 32; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
Qy 5 TCGCCAGGGCCCAAGGAGCGCTTCGGCGACTACGTGGACCGCTTCTTCAA 56
Db 5 TCAGGAAGGGACCTAAAGAACCTTTTAGAGATTATGTAGATAGATTATATAA 56
RESULT 7
US-08-303-275-193/c
Sequence 193, Application US/08303275
Patent No. 5766598
GENERAL INFORMATION:
APPLICANT: Paoletti, Enzo
APPLICANT: Tartaglia, James
APPLICANT: Cox, William I.
TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS RECOMBINANT
TITLE OF INVENTION: POXVIRUS VACCINE
NUMBER OF SEQUENCES: 205
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford
ADDRESSEE: c/o William S. Frommer
STREET: 530 Fifth Avenue


```

? FILING DATE: 29-OCT-1987
?
? PRIOR APPLICATION DATA:
?
? APPLICATION NUMBER: US 06/927,258
?
? FILING DATE: 04-NOV-1986
?
? ATTORNEY/AGENT INFORMATION:
?
? NAME: Trecartin, Richard F.
?
? REGISTRATION NUMBER: 311601
?
? REFERENCE/DOCKET NUMBER: A-55186-10/WHO
?
? TELECOMMUNICATION INFORMATION:
?
? TELEPHONE: (415) 781-1989
?
? TELEFAX: (415) 398-3249
?
? TELEX: 910 277299
?
? INFORMATION FOR SEQ ID NO: 47:
?
? SEQUENCE CHARACTERISTICS:
?
? LENGTH: 72 base pairs
?
? TYPE: nucleic acid
?
? STRANDEDNESS: unknown
?
? TOPOLOGY: unknown
?
? MOLECULE TYPE: DNA (genomic)
?
? US-08-707-237A-47

```

Query Match	30.0%	Score 18;	DB 2;	Length 72;
Best Local Similarity	100.0%;	Pred. No. 2.1e-03;		
Matches 18;	Conservative 0;	Mismatches 0;	Indels	

QY 2 ACA TCCG CCGAG GGGCCCCA 19
 | . | | | | : | | | | |
DB 26 ACA TCCG CCGAG GGGCCCCA 43

```

RESULT 1:
US-07-797-556-7/c
; Sequence 7, Application JS/57797556
; Patent No. 5262522
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; TITLE OF INVENTION: Receptor for Oncostatin M and Leukemia
; TITLE OF INVENTION: Inhibitory Factor
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/07/797,556
; FILING DATE: 19911122
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Seese, Kathryn A.
; REGISTRATION NUMBER: 32,172
; REFERENCE/DOCKET NUMBER: 2607
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-587-0430
; TELEFAX: 206-587-0606
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 100 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-07-797-556-7

```

Query Match 30.0%; Score 18; DB 1; Length 100;
Best Local Similarity 56.9%; Pred. No. 2.ic+03;

	Matches	33;	Conservative	0;	Mismatches	25;	Indels	0;	Gaps	0;
QY	3	CATCGGCACGGCCCAAGAGCCCTTCGCGAGCTACGTGGACCGCTTCTTCAAGACC	60							
DB	73	CCCTCCACCTGATCCACCACTCTCTGTACTCTCCACCACTGAACCCACCTCCACCGAAGAC	16							

RESULT 12
US-08-495-743-25/C
; Sequence 25, Application US/08495743
; Patent No. 5654416
; GENERAL INFORMATION:
; APPLICANT: Thomas J. Cummins, Susan M. Atwood
; APPLICANT: Lynn Bergmeyer, John B. Findlay
; APPLICANT: John W.H. Sutherland, JoAnne H. Kerschner
; TITLE OF INVENTION: DIAGNOSTIC COMPOSITIONS, ELEMENTS,
; TITLE OF INVENTION: METHODS AND TEST KITS FOR
; TITLE OF INVENTION: AMPLIFICATION AND DETECTION OF TWO
; TITLE OF INVENTION: OR MORE TARGET DNA'S USING PRIMERS
; TITLE OF INVENTION: HAVING MATCHED MELTING TEMPERATURES
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eastman Kodak Company, Patent Legal Staff
; STREET: 343 State Street
; CITY: Rochester
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 14650 - 2201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5inch, 1.44 Mb storage (IBM)
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS Version 3.3
; SOFTWARE: PC-8 (Word for Windows)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/495,743
; FILING DATE: 08-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/062,023
; FILING DATE: 14-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Tucker, J. Lanny
; REGISTRATION NUMBER: 27,678
; REFERENCE/DOCKET NUMBER: 67271A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 722-9332
; TELEFAX: (716) 477-4646
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 nucleotides
; TYPE: Nucleic acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; MOLECULE TYPE: Primer for HIV-1 DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE: Synthetically prepared
; IMMEDIATE SOURCE: Same
; PUBLICATION INFORMATION: Unknown
US-08-495-743-25

Query Match 29.7% Score 17.8; DB 1; Length 40;
 Best Local Similarity 75.9% Pred. NO. 2.3e+03;
 Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

07
1 GACATCCGCCAGGCGCCCCAAGGAGCCCTT 29
.IIII I II I I I IIII IIII

08
32 GACATAAGACAAGGACCACCAAGGAACCCCTT 4

RESULT 13
US-08-548-078-3/c
; Sequence 3, Application US/08548078

```
; Patent No. 5674717
; GENERAL INFORMATION:
; APPLICANT: John W. Backus
; APPLICANT: William H. Cornish
; APPLICANT: John B. Findlay
; APPLICANT: John W. H. Sutherland
; APPLICANT: Marlene M. King
; TITLE OF INVENTION: RAPID METHOD FOR PREFERENTIAL
; TITLE OF INVENTION: COMPLIFICATION OF TWO DIFFERENT NUCLEIC ACID
; TITLE OF INVENTION: SEQUENCES USING POLYMERASE CHAIN REACTION
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eastman Kodak Company,
; ADDRESSEE: Patent Legal Staff
; STREET: 343 State Street
; CITY: Rochester
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 14650 - 2201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 MB storage (IBM)
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS version 3.3
; SOFTWARE: PC-8 (word for windows 2.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/548,078
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA: No. 5674717e
; APPLICATION NUMBER: US/08/371,046
; FILING DATE:
; APPLICATION NUMBER: US/08/089,987
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Tucker, J. Lanny
; REGISTRATION NUMBER: 27,678
; REFERENCE/DOCKET NUMBER: 58258
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 722-9332
; TELEFAX: (716) 477-4646
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 nucleotides
; TYPE: Nucleic acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; MOLECULE TYPE: HIV-1 DNA primer
; HYPOTHETICAL: No
; ANTI-SENSE: No
; ORIGINAL SOURCE: Synthetically prepared
; IMMEDIATE SOURCE: Same
; PUBLICATION INFORMATION: No. 5674717e
; US-08-548-078-3

Query Match 29.7%; Score 17.8; DB 1; Length 40;
Best Local Similarity 75.9%; Pred. No. 2.3e+03;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 GACATCGCCGAGGCGCCCAAGGAGCCCTT 29
   |||| | || | | | | | | | | |
DB 32 GACATAAGACAGGACCAAGGAAACCCCTT 4

RESULT 14
US-08-495-739-25/c
; Sequence 25, Application US/08495739
; Patent No. 5702901
; GENERAL INFORMATION:
; APPLICANT: Thomas J. Cummins, Susan M. Atwood
; APPLICANT: Lynn Bergmeyer, John B. Findlay
; APPLICANT: John W.H. Sutherland, JoAnne H. Kerschner
; TITLE OF INVENTION: DIAGNOSTIC COMPOSITIONS, ELEMENTS,
; METHODS AND TEST KITS FOR
; AMPLIFICATION AND DETECTION OF TWO
; OR MORE TARGET DNA'S USING PRIMERS
; HAVING MATCHED MELTING TEMPERATURES
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eastman Kodak Company, Patent Legal Staff
; STREET: 343 State Street
; CITY: Rochester
; STATE: New York
; COUNTRY: U.S.A.
```

```
; TITLE OF INVENTION: AMPLIFICATION AND DETECTION OF TWO
; OR MORE TARGET DNA'S USING PRIMERS
; HAVING MATCHED MELTING TEMPERATURES
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eastman Kodak Company, Patent Legal Staff
; STREET: 343 State Street
; CITY: Rochester
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 14650 - 2201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 MB storage (IBM)
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS version 3.3
; SOFTWARE: PC-8 (word for windows)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/495,739
; FILING DATE: 08-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/062,023
; FILING DATE: 14 MAY 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Tucker, J. Lanny
; REGISTRATION NUMBER: 27,678
; REFERENCE/DOCKET NUMBER: 67271A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 722-9332
; TELEFAX: (716) 477-4646
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 nucleotides
; TYPE: Nucleic acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; MOLECULE TYPE: Primer for HIV-1 DNA
; HYPOTHETICAL: No
; ANTI-SENSE: No
; ORIGINAL SOURCE: Synthetically prepared
; IMMEDIATE SOURCE: Same
; PUBLICATION INFORMATION: Unknown
; US-08-495-739-25

Query Match 29.7%; Score 17.8; DB 1; Length 40;
Best Local Similarity 75.9%; Pred. No. 2.3e+03;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 GACATCGCCGAGGCGCCCAAGGAGCCCTT 29
   |||| | || | | | | | | | | |
DB 32 GACATAAGACAGGACCAAGGAAACCCCTT 4

RESULT 15
US-08-495-741-25/c
; Sequence 25, Application US/08495741
; Patent No. 5733751
; GENERAL INFORMATION:
; APPLICANT: Thomas J. Cummins, Susan M. Atwood
; APPLICANT: Lynn Bergmeyer, John B. Findlay
; APPLICANT: John W.H. Sutherland, JoAnne H. Kerschner
; TITLE OF INVENTION: DIAGNOSTIC COMPOSITIONS, ELEMENTS,
; METHODS AND TEST KITS FOR
; AMPLIFICATION AND DETECTION OF TWO
; OR MORE TARGET DNA'S USING PRIMERS
; HAVING MATCHED MELTING TEMPERATURES
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eastman Kodak Company, Patent Legal Staff
; STREET: 343 State Street
; CITY: Rochester
; STATE: New York
; COUNTRY: U.S.A.
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen, Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2003, 14:02:05 ; Search time 1380 Seconds
(without alignments)
1778.679 Million cell updates/sec

Title: US-09-475-704A-2
Perfect score: 60
Sequence: 1 gacatccgcagggcccaaa.....tggacccgtctcttcacagacc 60

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 1520254

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl :

- 1: gb_ba:*
- 2: gb_hgt:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: cm_ba:*
- 16: cm_fun:*
- 17: cm_hum:*
- 18: cm_in:*
- 19: cm_mu:*
- 20: cm_om:*
- 21: cm_or:*
- 22: cm_ov:*
- 23: cm_pat:*
- 24: cm_ph:*
- 25: cm_pl:*
- 26: cm_ro:*
- 27: cm_sts:*
- 28: cm_un:*
- 29: cm_vi:*
- 30: cm_htg_hum:*
- 31: cm_htq_inv:*
- 32: cm_htg_other:*
- 33: cm_htq_mus:*
- 34: cm_htg_pln:*
- 35: cm_htg_rtd:*
- 36: cm_htg_mam:*
- 37: cm_htg_vrt:*
- 38: cm_sy:*
- 39: cm_htgo_hum:*
- 40: cm_htgo_mus:*
- 41: cm_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	60	100.0	60	6	AX455886	AX455886 Sequence
2	55.2	92.0	60	6	AX455885	AX455885 Sequence
3	55.2	92.0	60	6	AX455937	AX455937 Sequence
4	34.4	57.3	60	6	AX455910	AX455910 Sequence
5	32.8	54.7	60	6	AX455938	AX455938 Sequence
6	21.9	36.3	84	6	AR021158	AR021158 Sequence
7	21.4	35.7	97	6	AF223423	AF223423 Pseudomon
8	21.2	35.3	84	9	HSU91013	U91013 Homo sapien
9	20.8	34.7	84	6	ARC21157	ARC21157 Sequence
10	20.8	34.7	84	6	AR021159	AR021159 Sequence
11	20.4	34.0	60	6	AR169921	AR169921 Sequence
12	20	33.3	81	6	AX480659	AX480659 Sequence
13	19.6	32.7	96	6	AX088790	AX088790 Sequence
14	19.4	32.3	51	9	HSDNASFAT	222386 H.sapiens D
15	19.4	32.3	53	6	I08503	I08503 Sequence 17
16	19.4	32.3	65	6	I08502	I08502 Sequence 16
17	19.4	32.3	78	6	AX480660	AX480660 Sequence
18	19.2	32.0	51	6	AX158488	AX158488 Sequence
19	19.2	32.0	100	9	AJ8116	AJ008116 Homo sapi
20	18.8	31.3	30	6	AX026570	AX026570 Sequence
21	19.8	31.3	74	5	AY227159	AY227159 Alca tord
22	18.6	31.0	51	6	AX157786	AX157786 Sequence
23	18.6	31.0	60	6	AX701057	AX701057 Sequence
24	18.4	30.7	51	6	AX203926	AX203926 Sequence
25	18.4	30.7	62	6	A10076	A10076 Epstein-Bar
26	18.4	30.7	62	14	EBVCLGAG	XC7531 EBV B95-8 C
27	18.4	30.7	69	10	AY205871	AY205871 Mus muscu
28	18.2	30.3	88	6	AX146316	AX146316 Sequence
29	18.2	30.3	99	6	AX523879	AX523879 Sequence
30	18	30.0	38	6	AR164537	AR164537 Sequence
31	18	30.0	72	6	AR051746	AR051746 Sequence
32	18	30.0	72	9	HSU91219	U91219 Homo sapien
33	18	30.0	80	14	AF040847	AF040847 Hepatitis
34	18	30.0	80	14	AF040849	AF040849 Hepatitis
35	18	30.0	98	3	ASBPB23A2	AF134278 Agrotis s
36	18	30.0	98	3	ASBPB23B2	AF134280 Agrotis s
37	18	30.0	98	3	ASBPB24A2	AF134282 Agrotis s
38	18	30.0	98	3	ASBPB25A2	AF134284 Agrotis s
39	18	30.0	98	3	ASBPB26A2	AF134286 Agrotis s
40	18	30.0	98	3	ASBPB27A2	AF134288 Agrotis s
41	17.8	29.7	40	6	AR142031	AR142031 Sequence
42	17.8	29.7	40	6	I59898	I59898 Sequence 25
43	17.8	29.7	40	6	I68065	I68065 Sequence 3
44	17.8	29.7	40	6	I86756	I86756 Sequence 25
45	17.8	29.7	40	6	I95781	I95781 Sequence 25

ALIGNMENTS

RESULT 1	AX455886	Sequence 2 from Patent WO0204493.	60 bp	DNA	linear	PAT 06-JUL-2002
LOCUS	AX455886	Sequence 2 from Patent WO0204493.	60 bp	DNA	linear	PAT 06-JUL-2002
DEFINITION	AX455886	Sequence 2 from Patent WO0204493.	60 bp	DNA	linear	PAT 06-JUL-2002
ACCESSION	AX455886	Sequence 2 from Patent WO0204493.	60 bp	DNA	linear	PAT 06-JUL-2002
VERSION	AX455886.1	GI:21714879	60 bp	DNA	linear	PAT 06-JUL-2002
KEYWORDS	AX455886.1	GI:21714879	60 bp	DNA	linear	PAT 06-JUL-2002
SOURCE	AX455886.1	GI:21714879	60 bp	DNA	linear	PAT 06-JUL-2002
ORGANISM	AX455886.1	GI:21714879	60 bp	DNA	linear	PAT 06-JUL-2002
REFERENCE	AX455886.1	GI:21714879	60 bp	DNA	linear	PAT 06-JUL-2002
AUTHORS	AX455886.1	GI:21714879	60 bp	DNA	linear	PAT 06-JUL-2002
TITLE	AX455886.1	GI:21714879	60 bp	DNA	linear	PAT 06-JUL-2002
1	Human immunodeficiency virus					
2	Human immunodeficiency virus					
3	Viruses; Retroviridae; Retroviridae; Lentivirus; Primate					
4	lentivirus group.					
5	zur Megede, J., Barnett, S.W., Engelbrecht, S. and van Rensburg, E.					
6	polynucleotides encoding antigenic hiv type c polypeptides,					
7	polypeptides and uses thereof					

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2003, 14:02:05 : Search time 193 Seconds
(without alignments)
839.204 Million cell updates/sec

Title: US-09-475-704A-2

Perfect score: 60

Sequence: 1 qacatccgcacgggccccaa.....tggaccgcttattcaagacc 60

Scoring table: IDENTITY_NJC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 2722628

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03:*
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*
25: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	60	100.0	60	21	AAA51608 HIV Gag major homo
2	58.4	97.3	60	21	AAA70426 HIV p55 Gag Major
3	55.2	92.0	60	21	AAA51607 HIV Gag major homo
4	55.2	92.0	60	24	ABL39974 Synthetic Gag poly
5	34.4	57.3	60	21	AAA51627 Major homology reg
6	33	55.0	90	24	ABK36257 HIV DNA encoding G
7	32.8	54.7	60	24	ABL39975 Wild type 8_5_TVL
8	31.8	53.0	87	7	AAN60663 HIV virus gag regi

9	29.8	49.7	90	24	ABK36256 HIV DNA encoding G
10	21.8	36.3	84	17	AAT33692 Nucleic acid encod
11	21	35.0	87	22	ABA71762 Human foetal liver
12	21	35.0	87	22	AAK20151 Human brain expres
13	21	35.0	87	22	AAK46211 Human bone marrow
14	21	35.0	87	22	AAI52117 Probe #20803 used
15	21	35.0	87	23	ABS45930 Human liver singlc
16	21	35.0	87	24	ABS20522 Human genome-deriv
17	20.8	34.7	84	17	AAT33691 Nucleic acid encod
18	20.8	34.7	84	17	AAT33693 Nucleic acid encod
19	20.4	34.0	60	24	AAD32141 CTLB36.1 antigen p
20	20.4	34.0	65	24	ABN52351 Mouse spliced tran
21	20.2	33.7	65	24	ABN57741 Loop 1b oligonucle
22	20	33.3	81	24	ABS66623 Codon-optimised HP
23	19.6	32.7	96	22	AAF75498 HIV reverse transc
24	19.4	32.3	45	22	AAD09440 Basic fibroblast g
25	19.4	32.3	53	18	AAT71251 Basic fibroblast g
26	19.4	32.3	65	18	AA771250 Loop 1c oligonucle
27	19.4	32.3	78	24	ABS66624 Novel murine polyn
28	19.4	32.3	100	24	ABS69273 Human silent SNP c
29	19.2	32.0	51	22	AAI74875 Primer R2 used to
30	18.8	31.3	30	21	AAA46684 Human calmodulin-1
31	18.8	31.3	51	21	AAA76375 Human silent SNP c
32	18.6	31.0	51	22	AAI74173 Human spliced tran
33	18.6	31.0	60	24	ABN36996 Sequence of primer
34	18.4	30.7	37	13	AAO21583 Human zinc finger
35	18.4	30.7	41	24	ABA04315 Human SNP oligonuc
36	18.4	30.7	50	22	AAL34430 Human SNP oligonuc
37	18.4	30.7	51	22	AAL32838 Human DNA containi
38	18.4	30.7	51	22	AAH79417 Human spliced tran
39	18.4	30.7	60	24	ABN40591 DNA sequence which
40	18.4	30.7	62	10	AAN91755 Topoisomerase vect
41	18.2	30.3	52	24	ABK8871 Human spliced tran
42	18.2	30.3	60	24	ABN35235 West Nile virus Ca
43	18.2	30.3	89	24	AAL44429 B-domain-deleted-F
44	18.2	30.3	99	24	ABS68080 infectious pancrea
45	18	30.0	38	20	AAZ25414

ALIGNMENTS

RESULT 1
AAA51608
10 AAA51608 standard; DNA; 60 BP.
XX AAA51608;
AC AAA51608;
XX 31-OCT-2000 (first entry)
DT HIV Gag major homology region nucleotides 841-900.
XX HIV Gag major homology region nucleotides 841-900.
DE Gag: expression cassette; antigenic; type C; HIV; Env: synthetic;
KW DNA immunization; packaging cell line; antigen presentation; ss.
KW Human immunodeficiency virus.
XX WO200039304-A2.
XX 06-JUL-2000.
XX 30-DEC-1999; 99WO-US31273.
XX 31-DEC-1998; 98US-0114495.
PR 01-SEP-1999; 99US-0152195.
XX (CHIR) CHIRON CORP.
PA Barnett S, Zur Megede J;
PI WPI; 2000-452401/39.
XX Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV

PT Env polypeptide and the polypeptide useful for immunizing a mammal:
XX especially human against HIV
PS Claim 1: Page 92; 113pp; English.
XX
CC Expression cassettes comprising a polynucleotide encoding antigenic
CC type C human immunodeficiency virus (HIV) Gag or Env polypeptides are
CC useful in DNA immunization, generation of packaging cell lines and
CC production of Gag- and/or Env-containing proteins. Synthetic Env and Gag
CC expression cassettes exhibit increased potency for induction of
CC cytotoxic T-lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1
CC self-assemble into non-infectious virus-like particles which are used as
CC a matrix for the proper presentation of an antigen entrapped or
XX associated to the immune system of the host.
SQ Sequence 60 BP; 12 A; 24 C; 15 G; 9 T; 0 other;

Query Match 100.0%; Score 60; DB 21; Length 60;
Best Local Similarity 100.0%; Pred. No. 3e-09;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATCCGCCAGGGCCCCCAGGAGGCCCTCCGCGACTACGTGGACCGCTTCTCAAGACC 60
DB 1 GACATCCGCCAGGGCCCCCAGGAGGCCCTCCGCGACTACGTGGACCGCTTCTCAAGACC 60

RESULT 2
AAA70426
ID AAA70426 standard; DNA: 60 BP.
XX
AC AAA70426;
XX
XX 28-NOV-2000 (first entry)
XX HIV p55 Gag Major Homology Region coding sequence.
DE HIV-1; AIDS; Gag; vaccine; expression cassette; ss.
KW Human immunodeficiency virus type 1.
XX
XX
XX
FH Key location/Qualifiers
FT CDS 1..60
FT /*tag= a
FT /partial
FT /product= "HIV p55 Gag Major Homology Region protein"
FT /note= "No stop codon given"
XX
XX WO200039302-A2.
XX
XX 06-JUL-2000.
XX
XX 30-DEC-1999; 99WO-US31245.
XX
XX 31-DEC-1998; 98US-0114495.
PR 01-DEC-1999; 99US-0168471.
XX
XX (CHIR) CHIRON CORP.
XX Barnett S. Zur Megede J. Srinivastava I. Lian Y. Hartog K. Liu H.
PI Greer C. Selby M. Walker C.
XX
XX WPI; 2000-452400/39.
DR P-PSDB; AAB14214.
XX
XX Expression cassettes encoding the human immunodeficiency virus (HIV)
PT Gag-containing polypeptide useful for vaccinating against HIV
PT Infections and acquired immunodeficiency syndrome (AIDS) -
XX
XX Claim 1: Page 345; 391pp; English.
PS The present invention relates to synthetic HIV Gag expression cassettes.
XX The Gag protein of HIV is needed for the assembly of virus-like
CC particles. In addition, the Gag protein is involved in many stages of the

CC HIV life cycle, including assembly, virion maturation after particle
CC release and early post-entry steps in viral replication. The expression
CC cassette may be used for the recombinant expression of HIV
CC Gag-polypeptides which may then be used to vaccinate against HIV
CC infection and acquired immunodeficiency syndrome (AIDS). The present
CC sequence is the coding sequence of HIV p55 Gag Major Homology Region.
CC This sequence is located within the p24-ca sequence of Gag. Mutations in
CC this region can severely impair particle production.
XX
SQ Sequence 60 BP; 13 A; 24 C; 15 G; 8 T; 0 other;

Query Match 97.3%; Score 58.4; DB 21; Length 60;
Best Local Similarity 98.3%; Pred. No. 9.3e-09;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACATCCGCCAGGGCCCCCAGGAGGCCCTCCGCGACTACGTGGACCGCTTCTCAAGACC 60
DB 1 GACATCCGCCAGGGCCCCCAGGAGGCCCTCCGCGACTACGTGGACCGCTTCTCAAGACC 60

RESULT 3
AAA51607
ID AAA51607 standard; DNA: 60 BP.
XX
AC AAA51607;
XX
XX 31-OCT-2000 (first entry)
XX HIV Gag major homology region nucleotides 844-903.
DE
XX
XX Gag; expression cassette; antigenic; type C; HIV; Env; synthetic;
KW DNA immunization; packaging cell line; antigen presentation; ss.
XX
XX Human immunodeficiency virus.
XX WO200039304-A2.
XX 06-JUL-2000.
XX
XX 30-DEC-1999; 99WO-US31273.
XX
XX 31-DEC-1998; 98US-0114495.
PR 01-SEP-1999; 99US-0152195.
XX
XX (CHIR) CHIRON CORP.
XX Barnett S. Zur Megede J.
XX WPI; 2000-452401/39.
XX
XX Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV
XX Env polypeptide and the polypeptide useful for immunizing a mammal
XX especially human against HIV
XX Claim 1: Page 92; 113pp; English.
XX
XX Expression cassettes comprising a polynucleotide encoding antigenic
XX type C human immunodeficiency virus (HIV) Gag or Env polypeptides are
XX useful in DNA immunization, generation of packaging cell lines and
XX production of Gag- and/or Env-containing proteins. Synthetic Env and Gag
XX expression cassettes exhibit increased potency for induction of
XX cytotoxic T-lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1
XX self-assemble into non-infectious virus-like particles which are used as
XX a matrix for the proper presentation of an antigen entrapped or
XX associated to the immune system of the host.
SQ Sequence 60 BP; 14 A; 22 C; 15 G; 9 T; 0 other;

Query Match 92.0%; Score 55.2; DB 21; Length 60;
Best Local Similarity 95.0%; Pred. No. 8.7e-08;
Matches 57; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GACATCCGCCAGGGCCCCCAGGAGGCCCTCCGCGACTACGTGGACCGCTTCTCAAGACC 60

Db 86 AGGAGCCCATCCCAACACTACCTGACACCATGATCAA 50
||||| ||| |||| || ||| || ||| ||||
RESULT 13
AAK46211/c
ID AAK46211 standard; DNA; 87 BP.
XX
AC AAK46211;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed single exon probe SEQ ID NO: 20768.
XX
XX Human; bone marrow expressed exon; gene expression analysis: probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00668.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
analyzing gene expression in human bone marrow.
XX
PS Example 4; SEQ ID NO: 20768; 658pp - Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
probes which are derived from genomic sequences expressed in the human
bone marrow. They can be used to measure gene expression in bone marrow
samples, which may enable the improved diagnosis and treatment of cancers
such as lymphoma, leukaemia and myeloma. The present sequence is one of
the probes of the invention.
XX
SQ Sequence 87 BP; 17 A; 14 C; 29 G; 27 T; 0 other;
Query Match 35.0%; Score 21; DB 22; Length 87;
Best Local Similarity 73.0%; Pred. No. 2.1e+03;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 20 AGGAGCCCTTCCGGACACTACCTGGACCGCTTCTTCAA 56
||||| ||| |||| || ||| || ||| ||||
Db 85 AGGAGCCCAICCAACACTACCTGAACCATGATCAA 50
RESULT 14
AAI52117/c
ID AAI52117 standard; DNA; 87 BP.
XX
AC AAI52117;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #20803 used to measure gene expression in human placenta sample.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;

genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00663.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-48897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
analyzing gene expression in human placenta.
XX
PS Claim 25; SEQ ID No 20803; 654pp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENP).
The present sequence is one such probe. The probes are useful for
producing a microarray for predicting, measuring and displaying gene
expression in samples derived from human placenta. The probes are useful
for antenatal diagnosis of human genetic disorders.
XX
SQ Sequence 87 BP; 17 A; 14 C; 29 G; 27 T; 0 other;
Query Match 35.0%; Score 21; DB 22; Length 87;
Best Local Similarity 73.0%; Pred. No. 2.1e+03;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 20 AGGAGCCCTTCCGGACACTACCTGGACCGCTTCTTCAA 56
||||| ||| |||| || ||| || ||| ||||
Db 86 AGGAGCCCATCCCAACACTACCTGAACCATGATCAA 50
RESULT 15
ABS45930/c
ID ABS45930 standard; DNA; 87 BP.
XX
AC ABS45930;
XX
DT 25-FEB-2003 (first entry)
XX
DE Human liver single exon probe, SEQ ID No 20920.
XX
KW Human; single exon nucleic acid probe; liver; cirrhosis;
KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW coronary heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157273-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00664.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488898/53.
XX Human genome-derived single exon nucleic acid probes useful for
PT analysing gene expression in human adult liver -
XX
PS Claim 4: SEQ ID NO 20920; 658bp; English.
XX
CC The invention relates to a single exon nucleic acid probe (SENP) (1) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13103 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult
CC liver. (1) may be used for predicting, measuring and displaying gene
CC expression in samples derived from human adult liver. The genes
CC identified may be involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which
CC is associated with coronary heart disease. ABS25011-ABS51005 represent
CC human liver single exon nucleic acid probes of the invention.
CC Note: The sequence information for this patent does not appear in the
CC printed specification but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 87 BP; 17 A; 14 C; 29 G; 27 T; 0 other;

Query Match 35.0%; Score 21; DB 23; Length 87;
Best Local Similarity 73.0%; Pred. No. 2.1e+03;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 20 AGGAGCCCTTCGGGACTACGTGGACGGCTTCTCAA 56
||||||| ||| ||||| || ||| ||| |||
Db 86 AGGAGCCCATCCNACACTACCTGAACACATGATCAA 50

Search completed: September 22, 2003, 14:08:51
Job time : 195 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compuen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2003, 14:02:05 : Search time 1798.5 seconds
(without alignments)
810.825 Million cell updates/sec

Title: US-09-475-704A-2
Perfect score: 60
Sequence: 1 gacatccgccagggccccc.....tggaccgcttcttcaagacc 60

Scoring table: IDENTITY_NJC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 452990

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22.2	37.0	88	28	BH413628
2	22.2	37.0	88	28	BH413629
3	21.8	36.3	98	28	AF039817
4	21.4	35.7	50	9	AU108059

5	21.4	35.7	64	14	H27186	H27186 y115j09.r1
6	21.4	35.7	93	14	W35927	W35927 mc32a07.r1
7	20.8	34.7	92	28	AZ920060	AZ920060 1006017H0
8	20.4	34.0	58	9	AI048494	AI048494 uh61a10.r
9	20.4	34.0	95	13	BQ760009	BQ760009 EBP107_SQ
10	20.2	33.7	74	9	AI272982	AI272982 qv63c07.x
11	20.2	33.7	77	9	AW022596	AW022596 df41b06.y
12	20.2	33.7	85	14	R12059	R12059 y54b11.r1
13	20.2	33.7	91	14	T48611	T48611 yb01a10.r1
14	20	33.3	75	29	CC043452	CC043452 3591_1_15
15	20	33.3	79	9	AA772825	AA772825 ae74e12.s
16	19.6	32.7	79	9	AA456718	AA456718 aa13h07.r
17	19.6	32.7	98	14	H49859	H49859 yo24e10.s1
18	19.4	32.3	50	9	AU102742	AU102742 AU102742
19	19.4	32.3	84	9	AA023110	AA023110 mh66c10.r
20	19.2	32.0	97	9	AA124117	AA124117 mp91h01.r
21	19	31.7	70	9	AI036945	AI036945 uh23b11.r
22	19	31.7	70	9	AI154612	AI154612 ud33d05.r
23	19	31.7	82	9	AA615572	AA615572 vo79d07.r
24	19	31.7	85	9	AA241182	AA241182 mv27e07.r
25	19	31.7	95	13	BU652348	BU652348 1112099D0
26	18.8	31.3	50	9	AU102354	AU102354 AU102354
27	18.8	31.3	55	12	BM285363	BM285363 EST00004
28	18.8	31.3	55	12	BM493218	BM493218 EST00001
29	18.8	31.3	55	13	BQ094077	BQ094077 040802_35
30	18.6	31.0	43	28	AZ417237	AZ417237 1M0192N18
31	18.6	31.0	74	29	CC249304	CC249304 RRF072_Ra
32	18.6	31.0	78	29	CC027848	CC027848 3591_1_7_
33	18.6	31.0	83	10	BE139376	BE139376 xr69g02.x
34	18.6	31.0	84	28	BH222409	BH222409 1006107B0
35	18.6	31.0	86	9	AA627888	AA627888 nq70a06.s
36	18.6	31.0	91	12	BG963271	BG963271 602828373
37	18.6	31.0	92	14	F30532	F30532 HSPD21049_!
38	18.6	31.0	95	29	CNS031GR	AL245556 Tetraodon
39	18.6	31.0	100	29	CC027859	CC027859 3591_1_7_
40	18.4	30.7	57	29	CNS020JW	AL206789 Tetraodon
41	18.4	30.7	69	28	BH408318	BH408318 1007054E1
42	18.4	30.7	74	29	CC043506	CC043506 3591_1_15
43	18.4	30.7	77	10	BF303006	BF303006 602030933
44	18.4	30.7	78	9	AW075659	AW075659 xb29f08.x
45	18.4	30.7	79	12	BI083018	BI083018 602874538

ALIGNMENTS

RESULT 1
BH413628
LOCUS BH413628 88 bp DNA linear GSS 12-DEC-2001
DEFINITION 1007033F04.iEL_x1 1007 - RescueMu Grid H Zea mays genomic, genomic survey sequence.
ACCESSION BH413628
VERSION BH413628.1 GI:17589985
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 88)
AUTHORS Walbot,V.
TITLE Maize genomic sequences found using engineered RescueMu transposon.
JOURNAL Unpublished
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Very probable ligation site of ends cut by single endonuclease.
Reverse complemented post-ligation sequence from source sequence.
Plate: 1007033 column: 6

```

Class: transposon-tagged.
FEATURES
  source
    Location/Qualifiers
      1..88
        /organism="Zea mays"
        /mol_type="genomic DNA"
        /cultiyar="mixed background W23/A188/B73"
        /db_xref="taxon:4577"
        /tissue_type="leaf"
        /dev_stage="adult"
        /lab_host="DH10B"
        /clone_lib="1007 - RescueMu Grid H"
        /note="Organ: leaf; Vector: RescueMu (engineered from
        pBlueScript backbone); Site_1: BamHI; Site_2: BglII;
        RescueMu is a 4.9 kb, modified maize Mu transposon
        designed to allow plasmid rescue from total genomic DNA.
        Mu elements insert preferentially into transcription
        units. For more information on RescueMu, go to the web
        site 'www.zmdb.iastate.edu' and follow the links for
        'RescueMu.' Grid H was grown at Berkeley in 2001. DNA
        was extracted from leaf punches, double digested using
        BamHI and BglII, and ligated to form circular plasmids.
        DH10B cells were transformed and then screened on LB
        plates with ampicillin."
  BASE COUNT 17 a 29 c 32 g 10 t
  ORIGIN

Query Match 37.0%; Score 22.2; DB 28; Length 88;
Best Local Similarity 64.7%; Pred. No. 5.5e+03;
Matches 33; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 GACATCCGCGAGGCCCCAAGGAGCCCTCCGGGACTACGTGGACCGCTTC 51
    ||||| 1 11 1 11 111 111 111 111 11 11
DB 38 GACATCTTCGCGAGGGGAGGCCCCCTCCGGGACTACGTGGCGGGGATC 88

RESULT 3
AF039817/c
LOCUS
DEFINITION
  AF039817 Human Homo sapiens genomic clone ET33.5, genomic survey
  sequence.
ACCESSION
  AF039817.1 GI:3253280
VERSION
  AF039817.1
KEYWORDS
  GSS.
SOURCE
  Homo sapiens (human)
  ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE
    1 (bases 1 to 98)
    Whitmore,S.A., Crawford,J., Apostolou,S., Eyre,H., Baker,E., Lower
    ,K.M., Settasatian,C., Goldup,S., Seshadri,R., Gibson,R.A., Mathew
    ,C.G., Cleton-Jansen,A.M., Savoia,A., Pronk,J.C., Auerbach,A.D.,
    Doggett,N.A., Sutherland,G.R. and Callen,D.F.
    Construction of a high-resolution physical and transcription map of
    chromosome 16q24.3: A region of frequent loss of heterozygosity in
    sporadic breast cancer
    Genomics 50 (1), 1-8 (1998)
    MEDLINE
    98292488
    PUBMED
    9628816
    COMMENT
      Contact: Whitmore SA
      Department of Cyto Genetics and Molecular Genetics
      Womens and Childrens Hospital
      72 King William Rd, North Adelaide, SA 5006, Australia
      Email: swhitmore@pulse.mad.adelaide.edu.au
      Class: unknown.
  FEATURES
    Location/Qualifiers
      1..98
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
        /map="16q24.3"
        /clone="ET33.5"
        /clone_lib="Human"
  BASE COUNT 31 a 21 c 34 g 12 t
  ORIGIN

Query Match 36.3%; Score 21.8; DB 28; Length 98;
Best Local Similarity 70.7%; Pred. No. 7.4e+03;
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 15 CCCCAGGAGCCCTCCGGGACTACGTGGACCGCTTCITCA 55
    111111111111111111111111111111111111
DB 77 CGCCTTGGAGTCTCTCTCCGCCTTCTTGCCAGCTTCTCCA 37

Class: transposon-tagged.
FEATURES
  source
    Location/Qualifiers
      1..88
        /organism="Zea mays"
        /mol_type="genomic DNA"
        /culti:var="mixed background W23/A188/B73"
        /db_xref="taxon:4577"
        /tissue_type="leaf"
        /dev_stage="adult"
        /lab_host="DH10B"
        /clone_lib="1007 - RescueMu Grid H"
        /note="Organ: leaf; Vector: RescueMu (engineered from
        pBlueScript backbone); Site_1: BamHI; Site_2: BglII;
        RescueMu is a 4.9 kb, modified maize Mu transposon
        designed to allow plasmid rescue from total genomic DNA.
        Mu elements insert preferentially into transcription
        units. For more information on RescueMu, go to the web
        site 'www.zmdb.iastate.edu' and follow the links for
        'RescueMu.' Grid H was grown at Berkeley in 2001. DNA
        was extracted from leaf punches, double digested using
        BamHI and BglII, and ligated to form circular plasmids.
        DH10B cells were transformed and then screened on LB
        plates with ampicillin."
  BASE COUNT 17 a 29 c 32 g 10 t
  ORIGIN

Query Match 37.0%; Score 22.2; DB 28; Length 88;
Best Local Similarity 64.7%; Pred. No. 5.5e+03;
Matches 33; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 GACATCCGCGAGGCCCCAAGGAGCCCTCCGGGACTACGTGGACCGCTTC 51
    ||||| 1 11 1 11 111 111 111 111 11 11
DB 38 GACATCTTCGCGAGGGGAGGCCCCCTCCGGGACTACGTGGCGGGGATC 88

RESULT 2
BH413629
LOCUS
DEFINITION
  BH413629 88 bp DNA linear GSS 12-DHC-2001
  1007033F04.y1 1007 - RescueMu Grid H Zea mays genomic, genomic
  survey sequence.
ACCESSION
  BH413629
VERSION
  BH413629.1 GI:17589989
KEYWORDS
  GSS.
SOURCE
  Zea mays
  ORGANISM
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
    clade; Panicoideae; Andropogoneae; Zea.
  REFERENCE
    1 (bases 1 to 88)
    Walbot,V.
    Maize genomic sequences found using engineered RescueMu transposon
    Unpublished
    Contact: Walbot V
    Department of Biological Sciences
    Stanford University
    855 California Ave, Palo Alto, CA 94304, USA
    Tel: 650 723 2227
    Fax: 650 725 8221
    Email: walbot@stanford.edu
    Very probable ligation site found so sequence was trimmed.
    Post-ligation sequence submitted separately.
    Plate: 1007033 column: 6
    Class: transposon-tagged.
  FEATURES
    Location/Qualifiers
      1..88
        /organism="Zea mays"
        /mol_type="genomic DNA"
        /culti:var="mixed background W23/A188/B73"
        /db_xref="taxon:4577"
        /tissue_type="leaf"
        /dev_stage="adult"
        /lab_host="DH10B"
        /clone_lib="1007 - RescueMu Grid H"

```


High quality sequence stop: 84.

FEATURES
Source

Location/Qualifiers
1..93
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:350196"
/dev_stage="19.5 dpc total fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares mouse p3MMF19.5"
/note="Vector: pT73D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'] TGITACCAATCTGAAGTGGGAGCGCGCAITTTTITTTTITTTT 3'; double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Fatima Bonaldo. RNA was kindly provided by Dr. Minoru Ko (Wayne State University)."

BASE COUNT 19 a 28 c 28 g 18 t

ORIGIN

Query Match 35.7%; Score 21.4; DB 14; Length 93;
Best Local Similarity 71.8%; Pred. No. 9.7e+03;
Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 19 AAGGAGCCCTCCGGGACTACGTGACCGCTTCTTCAAG 57

Db 44 AAGCGGIACCTCTGTGACTACTGTGACCGCTCTCCAG 82

RESULT 7

A2920060

LOCUS

DEFINITION A2920060 92 bp DNA linear GSS 17-DEC-2001
1006017H08.x1 1006 - RescueMu Grid G Zea mays genomic, genomic survey sequence.

ACCESSION

VERSION A2920060.1 GI:13390392

KEYWORDS

SOURCE GSS.

ORGANISM

Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 92)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Maize genomic sequences found using engineered RescueMu transposon
Unpublished

Contact: Walbot V.

Department of Biological Sciences

Stanford University

855 California Ave, Palo Alto, CA 94304, USA

Tel: 650 723 2227

Fax: 650 725 8221

Email: walbot@stanford.edu

Plate: 1006017 row: 37

Class: transposon-tagged.

FEATURES

Source

Location/Qualifiers
1..92
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="mixed background W23/A188/B73"
/db_xref="taxon:4577"
/tissue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"

/clone_lib="1006 - RescueMu Grid G"
/note="Organ: leaf; Vector: RescueMu (engineered from pBlueScript backbone); Site_1: BamHI; Site_2: BglII; RescueMu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA.

Mu elements insert preferentially into transcription units. For more information on RescueMu, go to the web site 'www.zmdh.iastate.edu' and follow the links for 'RescueMu.' Grid G was grown at Stanford in 2000. DNA was extracted from leaf punches, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."

BASE COUNT 21 a 37 c 26 g 8 t

ORIGIN

Query Match 34.7%; Score 20.8; DB 28; Length 92;
Best Local Similarity 64.6%; Pred. No. 1.5e+04;
Matches 31; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 1 GACATCCGCGAGGCGCCCAAGGAGCGCTTCCGCGACTACGTGGACCGC 48

Db 24 GACAACCTCTCCGCGACTACAAGAGGTCCGCGACTACGAGTCCCGC 71

RESULT 8

A1048494/c

LOCUS

DEFINITION A1048494 58 bp mRNA linear EST 08-JUL-1998
uh61a10.r1 Soares embryonic_stem_cell_MMES Mus musculus cDNA clone
IMAGE:1749882 5' similar to SW:3BFL_MOUSE P55194 SH3-BINDING
PROTEIN 3BP-1.1, mRNA sequence.

ACCESSION A1048494.1 GI:3296781

VERSION EST.

KEYWORDS Mus musculus (house mouse)

SOURCE Mus musculus

ORGANISM

Fukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 58)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
The WashU-HMI Mouse EST Project
Unpublished
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through JLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

TITLE

JOURNAL

COMMENT

FEATURES

Source

Location/Qualifiers
1..58
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:1749882"
/cell_type="embryonic stem cell"
/lab_host="DH10B"
/clone_lib="Soares_embryonic_stem_cell_MMES"
/note="Organ: bone marrow; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5']
TGTTACCAATCTGAG1GGGAGCGCCCATGCTTTTITTTTITTTTITTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through two rounds of normalization, and was

BASE COUNT 19 a 27 c 15 g 18 t
ORIGIN

Query Match. 33.3%; Score 20; DB 9; Length 79;
Best Local Similarity 65.9%; Pred. NO. 2.5e+04;
Matches 29; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 17 CCAAGGAGCCCTTCGGCGACTACGTTGGACGGCTTCTTCAAGACC 60
 |||..||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3 CCGAGGAGTTCTGCGCGCATCTACCTGCTCGGCGATCCCGACACCC 45

Search completed: September 22, 2003, 15:57:08
Job time : 1801.5 secs

GenCore version 5.1.6
Copyright (C) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2003, 14:10:43 : Search time 144.5 Seconds
(without alignments)
1021.419 Million cell updates/sec

Title: US-09-475-704A-1

Perfect score: 60

Sequence: 1 gacatcaagcaggcccaaa.....tggaacgcttcttaaacacg 60

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1660708 seqs, 1229959015 residues

Total number of hits satisfying chosen parameters: 1177748

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_NA:*

- 1: /cgn2_5/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_5/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_5/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_5/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_5/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_5/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_5/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_5/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_5/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_5/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_5/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_5/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_5/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_5/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_5/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_5/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 17: /cgn2_5/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	21	35.0	87	9	US-09-864-761-28254
2	20.4	34.0	60	9	US-09-007-093-7
3	19	31.7	60	12	US-09-908-975-16956
4	18.8	31.3	65	12	US-09-908-975-24532
C 5	18.6	31.3	66	9	US-09-815-242-1315
6	18.6	31.0	60	12	US-09-908-975-9744
7	18.6	31.0	65	12	US-09-908-975-25099
8	18.6	31.0	65	12	US-09-908-975-30489
9	18.4	30.7	70	14	US-10-060-036-424
C 10	18.2	30.3	60	12	US-09-908-975-22576
11	18	30.0	60	12	US-09-908-975-14069
C 12	18	30.0	80	9	US-09-864-761-18009
C 13	17.8	29.7	60	12	US-09-908-975-13021
14	17.8	29.7	60	12	US-09-908-975-13510
15	17.8	29.7	60	12	US-09-908-975-20072
16	17.8	29.7	63	9	US-09-811-672-4

17	17.8	29.7	63	12	US-10-161-660-4	Sequence 4, Appli
C 18	17.8	29.7	65	12	US-09-908-975-1045	Sequence 1045, Ap
C 19	17.8	29.7	100	10	US-09-728-445-336	Sequence 336, App
20	17.8	29.7	100	10	US-09-908-855-45	Sequence 45, Appl
C 21	17.6	29.3	50	12	US-10-322-360-119	Sequence 119, Appl
C 22	17.6	29.3	56	9	US-09-760-574-10	Sequence 10, Appl
C 23	17.6	29.3	56	11	US-09-766-442A-10	Sequence 10, Appl
24	17.6	29.3	60	12	US-09-908-975-8727	Sequence 8727, Ap
25	17.6	29.3	65	12	US-09-908-975-23836	Sequence 23836, A
26	17.6	29.3	65	12	US-09-908-975-24465	Sequence 24465, A
27	17.4	29.0	58	11	US-09-747-377-332	Sequence 332, App
C 28	17.4	29.0	58	14	US-10-105-613-332	Sequence 332, App
C 29	17.4	29.0	60	12	US-09-908-975-13049	Sequence 13049, A
C 30	17.4	29.0	60	12	US-09-908-975-18261	Sequence 18261, A
C 31	17.4	29.0	65	12	US-09-908-975-27199	Sequence 27199, A
C 32	17.4	29.0	65	12	US-09-908-975-27227	Sequence 27227, A
33	17.4	29.0	65	12	US-09-908-975-27703	Sequence 27703, A
34	17.4	29.0	65	12	US-09-908-975-28587	Sequence 28587, A
35	17.4	29.0	73	11	US-09-805-353-17	Sequence 17, Appl
C 36	17.4	29.0	77	11	US-09-805-353-16	Sequence 16, Appl
C 37	17.2	28.7	25	14	US-10-098-263B-83464	Sequence 83464, A
C 38	17.2	28.7	48	12	US-10-295-798-40	Sequence 40, Appl
39	17.2	28.7	60	12	US-09-908-975-13613	Sequence 13613, A
40	17.2	28.7	100	10	US-09-783-590-914	Sequence 914, App
41	17	28.3	60	12	US-09-908-975-18498	Sequence 18498, A
42	17	28.3	65	12	US-09-908-975-28077	Sequence 28077, A
43	17	28.3	70	11	US-09-951-061A-87	Sequence 87, Appl
C 44	17	28.3	70	11	US-09-951-061A-88	Sequence 88, Appl
45	17	28.3	76	9	US-09-062-104A-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1

US-09-864-761-28254/c
: Sequence 28254, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David K.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL I
: FILE REFERENCE: Aomic-X-1
: CURRENT APPLICATION NUMBER: US/09/864.761
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263.6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662

```
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCI/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCI/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 28254
; LENGTH: 87
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004870.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
; OTHER INFORMATION: SWISSPROT HIT: P75080, EVAIJE 5.60e+00
US-09-864-761-28254

Query Match 35.0%; Score 21; Db 9; Length 87;
Best Local Similarity 73.0%; Pred. No. 2.5e+02;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 20 AGGAGCCCTCCGGGACTACGTGGACCGCTTCTTCAA 56
||||||| ||| ||||| ||| ||| ||| ||| |||
Db 86 AGGAGCCCAACCAACACTACCTGAACCAATGATCAA 50

RESULT 2
US-09-007-093-7
; Sequence 7, Application US/09007093
; Patent No. US2002025315A1
; GENERAL INFORMATION:
; APPLICANT: Anand, Naveen N
; APPLICANT: Barber, Brian H
; APPLICANT: Cates, George A
; APPLICANT: Caterini, Judith F
; APPLICANT: Klein, Michel H
; TITLE OF INVENTION: CHIMERIC ANTIBODIES FOR DELIVERY OF
; TITLE OF INVENTION: ANTIGENS TO SELECTED CELLS OF THE IMMUNE SYSTEM
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/007,093
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/483,576
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael J
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-765

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 60 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-007-093-7

Query Match 34.0%; Score 20.4; Db 9; Length 60;
Best Local Similarity 65.2%; Pred. No. 4e+02;
Matches 30; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 13 GGCCCCAAGGAGCCCTTCGGGACTACGTGGACCGCTTCTTCAA 58
||||||| ||| ||||| ||| ||| ||| ||| ||| |||
Db 13 GGTCCTAAAGAACCTTTTAGAGACTATGTTGATAGGTTTATAAGA 58

RESULT 3
US-09-908-975-16956
; Sequence 16956, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPL
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16956
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-908-975-16956

Query Match 31.7%; Score 19; Db 12; Length 60;
Best Local Similarity 65.1%; Pred. No. 1.3e+03;
Matches 28; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 4 ATCAAGCAGGCCCCCAAGGAGCCCTTCGGGACTACGTGGACCGCTTCTTCAA 46
||||||| ||| ||||| ||| ||||| ||| ||| ||| |||
Db 5 AACGTGCGCGCTCCAAAGCGCTTTCGGGAGCCCTGCTGAGCC 47

RESULT 4
US-09-908-975-24532
; Sequence 24532, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPL
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
```

```
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 24532
; LENGTH: 65
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-908-975-24532

Query Match      31.3%; Score 18.8; DB 9; Length 65;
Best Local Similarity 59.3%; Pred. No. 1.5e+03;
Matches 32; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY      3 CAACAAGCAGGCGCCCAAGGAGCCCTTCGGCGACTACGTGGACCGCTTCTTCAAGAC 56
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      11 CCTAGAGCTGTGCCACGACGAACTCTTCGGCGGACCTCTTCAACAGCAACCAACA 64

RESULT 5
US-09-815-242-1315/C
; Sequence 1315, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in Prokaryotes
; FILE REFERENCE: ELITRA-011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,576
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1315
; LENGTH: 66
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-1315

Query Match      31.3%; Score 18.8; DB 9; Length 65;
Best Local Similarity 59.3%; Pred. No. 1.5e+03;
Matches 32; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY      6 CAACGAGGCGCCCAAGGAGCCCTTCGGCGACTACGTGGACCGCTTCTTCAAGAC 59
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      62 CAACCGCGGTTCCAAACCGCCCTTCCTGCCGCGGCTGCGGCAAGCC 9

RESULT 6
US-09-908-975-9744
; Sequence 9744, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Aion
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPI
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/257,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9744
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-908-975-9744

Query Match      31.0%; Score 18.6; DB 12; Length 60;
Best Local Similarity 65.9%; Pred. No. 1.8e+03;
Matches 27; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY      18 CAAGGAGCCCTTCGGCGACTACGTGGACCGCTTCTTCAAGA 58
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      7 CAAGGAGACAGTCGCGACGAGTTCATCTTCTACTCCAAGA 47

RESULT 7
US-09-908-975-25099
; Sequence 25099, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Aion
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPI
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 25099
; LENGTH: 65
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-908-975-25099

Query Match      31.0%; Score 18.6; DB 12; Length 65;
Best Local Similarity 61.2%; Pred. No. 1.8e+03;
Matches 30; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY      11 AGGGCCCCAAGGAGCCCTTCGGCGACTACGTGGACCGCTTCTTCAAGAC 59
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      7 AGACACCCATGGAGATGCTCTACCAACCGTGCACCGACTCAACATGTC 55

RESULT 8
US-09-908-975-30489
; Sequence 30489, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
```


RESULT 15

```

US-09-908-975-20072
; Sequence 20072, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Eilat
; APPLICANT: FAIGLER, Simchen
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
; FILE REFERENCE: 36686-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/267,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: Patent.In version 3.0
; SEQ ID NO 20072
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-908-975-20072

```

Query Match	29.7%	Score 17.8	DB 12	length 65
Best Local Similarity	62.2%	Pred. NO. 3.5e-03		
Matches	28	Conservative	0	Mismatches 17; Indels 0; Gaps 0;
QY	15	CCCCAAGGAGCGCTCCGGGAGTACGTGGCGCGCTCTTCGAAGAC	59	
D6	12	CCATCAGGAGTTCACGGCGACCAACTTCCAGCTCTTCATCAGAC	56	

Search completed: September 22, 2003, 17:39:57
Job time : 147.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

QM nucleic - nucleic search, using sw model

Run on: September 22, 2003, 14:02:05 : Search time 14 seconds
(without alignments)
602,886 Million cell updates/sec

Title: US-09-475-704A-1
Perfect score: 60
Sequence: 1 gacatcaagcagggccccc.....tggaacggttttcaagacc 60

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 830498

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA: *
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PCFUS_COMB.seq: *
6: /cgn2_5/ptodata/2/ina/backfiles.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21.6	36.0	72	1 US-08-303-275-192	Sequence 192, Appl
2	21.6	36.0	72	1 US-08-303-275-193	Sequence 193, Appl
3	20.4	34.0	60	3 US-08-943-136-7	Sequence 7, Appl
4	20.4	34.0	60	4 US-08-973-518-7	Sequence 7, Appl
5	19.6	32.7	57	6 5514566-17	Patent No. 5514566
6	18	30.0	100	1 US-08-009-265-33	Sequence 33, Appl
7	18	30.0	100	2 US-08-322-679-6	Sequence 6, Appl
8	18	30.0	100	4 US-08-686-993A-18	Sequence 18, Appl
9	17.8	29.7	40	1 US-08-495-743-25	Sequence 25, Appl
10	17.8	29.7	40	1 US-08-548-078-3	Sequence 3, Appl
11	17.8	29.7	40	1 US-08-495-739-25	Sequence 25, Appl
12	17.8	29.7	40	1 US-08-495-741-25	Sequence 25, Appl
13	17.8	29.7	40	3 US-08-062-023-25	Sequence 25, Appl
14	17.8	29.7	63	3 US-08-750-419A-4	Sequence 4, Appl
15	17.8	29.7	63	4 US-09-811-672-4	Sequence 4, Appl
16	17.8	29.7	65	3 US-09-191-852-16	Sequence 16, Appl
17	17.8	29.7	65	4 US-08-817-906-16	Sequence 16, Appl
18	17.8	29.7	65	5 PCT-US95-13376-16	Sequence 16, Appl
19	17.8	29.7	72	1 US-08-303-275-195	Sequence 195, Appl
20	17.8	29.7	100	3 US-09-242-690A-45	Sequence 45, Appl
21	17.6	29.3	50	4 US-09-554-929-119	Sequence 119, Appl
22	17.6	29.3	72	1 US-07-778-233B-72	Sequence 72, Appl
23	17.6	29.3	72	1 US-07-963-321-72	Sequence 72, Appl
24	17.6	29.3	72	1 US-08-290-641-72	Sequence 72, Appl
25	17.6	29.3	72	1 US-08-548-540-72	Sequence 72, Appl
26	17.6	29.3	72	5 PCT-US96-09809-72	Sequence 72, Appl
27	17.4	29.0	53	2 US-07-695-201B-13	Sequence 13, Appl

C	28	17.4	29.0	53	3	US-08-470-532-13	Sequence 13, Appl
	29	17.2	28.7	50	1	US-08-575-052-7	Sequence 7, Appl
	30	17.2	28.7	50	1	US-08-614-516A-7	Sequence 7, Appl
	31	17.2	28.7	50	1	US-08-770-557-7	Sequence 7, Appl
	32	17.2	28.7	50	1	US-08-405-234-7	Sequence 7, Appl
	33	17.2	28.7	63	4	US-09-455-679-3	Sequence 3, Appl
	34	17	28.3	70	1	US-08-105-483-406	Sequence 406, Appl
	35	17	28.3	70	1	US-08-224-657-89	Sequence 89, Appl
	36	17	28.3	70	1	US-08-224-657-90	Sequence 90, Appl
	37	17	28.3	70	1	US-08-709-209-406	Sequence 406, Appl
	38	17	28.3	70	1	US-08-257-073-44	Sequence 44, Appl
	39	17	28.3	70	1	US-08-257-073-45	Sequence 45, Appl
	40	17	28.3	70	1	US-08-303-275-112	Sequence 112, Appl
	41	17	28.3	70	1	US-08-458-101-406	Sequence 406, Appl
	42	17	28.3	70	2	US-08-658-665-129	Sequence 129, Appl
	43	17	28.3	70	2	US-08-658-665-130	Sequence 130, Appl
	44	17	28.3	70	3	US-08-675-566-71	Sequence 71, Appl
	45	17	28.3	70	3	US-08-675-566-72	Sequence 72, Appl

ALIGNMENTS

RESULT 1
US-08-303-275-192
Sequence 192, Application US/08303275
Patent No. 5766598
GENERAL INFORMATION:
APPLICANT: Paoletti, Enzo
APPLICANT: Tartaglia, James
APPLICANT: Cox, William I.
TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS RECOMBINANT
NUMBER OF SEQUENCES: 205
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Saiford
ADDRESSEE: C/O William S. Frommer
STREET: 530 Fifth Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/303,275
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/897,382
FILING DATE: 11-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2420
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 192:
SEQUENCE CHARACTERISTICS:
LENGTH: 72 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-303-275-192

Query Match 36.0%; Score 21.6; DB 1; Length 72;
Best Local Similarity 63.5%; Pred. No. 1.4e+02;
Matches 33; Conservative 0; Mismatches 19; Indels 0; Gaps 0;


```
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 2038-781
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-973-518-7

Query Match          34.0%; Score 20.4; DB 4; Length 60;
Best Local Similarity 65.2%; Pred. No. 3.2e+02;
Matches 30; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 13 GGCCCCAAGGAGCCCTCCGGGACTACGCTGGACCGCTTCTTCAAGA 58
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 13 GGTCTAAAGAACCTTTAGAGACTATGTTGATAGCTTTTATTAAGA 58

RESULT 5
5514566-17
PATENT NO. 5514566
APPLICANT: FIDDES, JOHN C.; ABRAHAM, JUDITH A.
TITLE OF INVENTION: METHODS OF PRODUCING RECOMBINANT
FIBROBLASTS GROWTH FACTORS
NUMBER OF SEQUENCES: 21
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/417,022
FILING DATE: 05-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 809,163
FILING DATE: 16-DEC-1985
APPLICATION NUMBER: 775,521
FILING DATE: 12-SEP-1985
SEQ ID NO: 17:
LENGTH: 57
5514566-17

Query Match          32.7%; Score 19.6; DB 6; Length 57;
Best Local Similarity 62.0%; Pred. No. 5.8e+02;
Matches 31; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1 AGGGCCCCAAGGAGCCCTCCGGGACTACGCTGGACCGCTTCTTCAAGACC 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2 AGGACCCCAAGGCGCTCTACTGCAAGAACGGGGCTTCTTCCGAGAAATC 51

RESULT 6
US-08-009-265-33
Sequence 33, Application US/08009265
Patent No. 5547871
GENERAL INFORMATION:
APPLICANT: Black Dr., Bruce C.
APPLICANT: Summers Dr., Max D.
TITLE OF INVENTION: Heterologous Signal Sequences For
TITLE OF INVENTION: Secretion Of Insect Controlling Proteins
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: American Cyanamid Company
STREET: 1937 West Main Street, P.O. Box 50
CITY: Stamford
STATE: CT
COUNTRY: USA
ZIP: 06904-0060
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/009,265
FILING DATE: 19930125
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gordon Mr., Alan M.
REGISTRATION NUMBER: 30637
REFERENCE/DOCKET NUMBER: 31868-00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 203-321-2719
TELEFAX: 203-321-2971
TELEX: 710-474-4059
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 100 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Scorpion Androctonus Australis Hector
US-08-009-265-33

Query Match          30.0%; Score 18; DB 1; Length 100;
Best Local Similarity 56.9%; Pred. No. 2e+03;
Matches 33; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 3 CATCAAGCAGGGCCCAAGSAGCCCTTCGCGACTACGCTGGACCGCTTCTTCAAGACC 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 31 CAACAATGAATGCACCAAGGTGCACCTACGCTGACAAAGGGCTACTGTGCTTCTGCTCC 98

RESULT 7
US-08-322-679-6
Sequence 6, Application US/08322679
Patent No. 5965123
GENERAL INFORMATION:
APPLICANT: Ahmed, Fakhruddin
TITLE OF INVENTION: Coated Pesticidal Agents, Processes For
TITLE OF INVENTION: Their Preparation And Compositions Containing Them
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: American Cyanamid Company
STREET: One Cyanamid Plaza
CITY: Wayne
STATE: New Jersey
COUNTRY: United States
ZIP: 07470
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/322,679
FILING DATE: 13-OCT-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Hogan, John W.
REGISTRATION NUMBER: 32,703
REFERENCE/DOCKET NUMBER: 32,458
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-831-3195
TELEFAX: 201-831-3305
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 100 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
```

```

; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-322-679-6

Query Match          30.0%; Score 18; DB 2; Length 100;
Best Local Similarity 56.9%; Pred. No. 2e+03;
Matches 33; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 3 CATCAAGCAGGCGCCCAAGAGCCCTTCCCGGACTACGTGGACCGCTTCTTCAAGACC 60
DB 31 CAACAATGAATGCACCAAGGTGCACTACGCTGACAGGCTACTGTGGCTTCTGTCAC 58

RESULT 8
US-08-686-993A-18
; Sequence 18, Application US/08686993A
; Patent No. 6355240
; GENERAL INFORMATION:
; APPLICANT: Dierks, Peter
; TITLE OF INVENTION: ENHANCED INSECTICIDAL INSECT VIRUS THROUGH THE
; TITLE OF INVENTION: EXPRESSION OF HETEROLOGOUS PROTEINS WITH EARLY
; TITLE OF INVENTION: PROMOTERS
; FILE REFERENCE: 0646/0D616
; CURRENT APPLICATION NUMBER: US/08/686,993A
; CURRENT FILING DATE: 1996-07-25
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 100
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: cDNA
US-08-686-993A-18

Query Match          30.0%; Score 18; DB 4; Length 100;
Best Local Similarity 56.9%; Pred. No. 2e+03;
Matches 33; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 3 CATCAAGCAGGCGCCCAAGAGCCCTTCCCGGACTACGTGGACCGCTTCTTCAAGACC 60
DB 31 CAACAATGAATGCACCAAGGTGCACTACGCTGACAGGCTACTGTGGCTTCTGTCAC 58

RESULT 9
US-08-495-743-25/c
; Sequence 25, Application US/08495743
; Patent No. 5654416
; GENERAL INFORMATION:
; APPLICANT: Thomas J. Cummins, Susan M. Atwood
; APPLICANT: Lynn Bergmeyer, John B. Findlay
; APPLICANT: John W.H. Sutherland, JoAnne H. Kerschner
; TITLE OF INVENTION: DIAGNOSTIC COMPOSITIONS, ELEMENTS,
; TITLE OF INVENTION: METHODS AND TEST KITS FOR
; TITLE OF INVENTION: AMPLIFICATION AND DETECTION OF TWO
; TITLE OF INVENTION: OR MORE TARGET DNA'S USING PRIMERS
; TITLE OF INVENTION: HAVING MATCHED MELTING TEMPERATURES
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eastman Kodak Company, Patent Legal Staff
; STREET: 343 State Street
; CITY: Rochester
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 14650 - 2201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5inch, 1.44 MB storage (IBM)
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS Version 3.3
; SOFTWARE: PC-8 (Word for Windows)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/495,743
```

```

; FILING DATE: 08-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/062,023
; FILING DATE: 14-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Tucker, J. Lanny
; REGISTRATION NUMBER: 27,678
; REFERENCE/DOCKET NUMBER: 67271A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 722-9332
; TELEFAX: (716) 477-4646
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 nucleotides
; TYPE: Nucleic acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; MOLECULE TYPE: Primer for HIV-1 DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE: Synthetically prepared
; IMMEDIATE SOURCE: Same
; PUBLICATION INFORMATION: Unknown
US-08-495-743-25

Query Match          29.7%; Score 17.8; DB 1; Length 40;
Best Local Similarity 75.9%; Pred. No. 2.1e+03;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 GACATCAAGCAGGCGCCCAAGAGCCCTT 29
DB 32 GACATAAGACACAGGACCACCAAGGAACCTT 4

RESULT 10
US-08-548-078-3/c
; Sequence 3, Application US/08548078
; Patent No. 5674717
; GENERAL INFORMATION:
; APPLICANT: John W. Backus
; APPLICANT: William H. Donish
; APPLICANT: John B. Findlay
; APPLICANT: John W. H. Sutherland
; APPLICANT: Marlene M. King
; TITLE OF INVENTION: RAPID METHOD FOR PREFERENTIAL
; TITLE OF INVENTION: COAMPLIFICATION OF TWO DIFFERENT NUCLEIC ACID
; TITLE OF INVENTION: SEQUENCES USING POLYMERASE CHAIN REACTION
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eastman Kodak Company,
; ADDRESSEE: Patent Legal Staff
; STREET: 343 State Street
; CITY: Rochester
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 14650 - 2201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 MB storage (IBM)
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS Version 3.3
; SOFTWARE: PC-8 (Word for Windows 2.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/548,078
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA: No. 5674717e
; APPLICATION NUMBER: US/08/371,046
; FILING DATE:
; APPLICATION NUMBER: US/08/089,987
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Tucker, J. Lanny
```


; TITLE OF INVENTION: RECOMBINANT ALLERGEN, FRAGMENTS THEREOF, CORRESPONDING RECOMBINANT
; TITLE OF INVENTION: MOLECULES, VECTORS AND HOSTS CONTAINING THE DNA MOLECULES, DIAGN
; TITLE OF INVENTION: THERAPEUTIC USES OF SAID ALLERGENS AND FRAGMENTS
; FILE REFERENCE: 1614-0247P
; CURRENT APPLICATION NUMBER: US/09/811,672
; CURRENT FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 63
; TYPE: DNA
; ORGANISM: Timothy Grass
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(63)
; OTHER INFORMATION:
CS-09-811-672-4

Query Match 29.7%; Score 17.8; DB 4; Length 63;
Best Local Similarity 58.5%; Pred. No. 2.2e+03;
Matches 31; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
Qy 2 ACATCAAGCAGGGCCCGAAGGAGGCGCTTCGGCGGACTACGTGGAGCGGCTTC 54
:||||| | | | | | | | | | | | | | | | | | | | |
Db 2 ACATCAGCGACGACAAAGGAGGAGGCGCATCGCCCGCTACACATTCGACCTCTCC 54
:||||| | | | | | | | | | | | | | | | | | | | |

Search completed: September 22, 2003, 14:10:33
Job time : 47 secs

GenCore version 5.1.6
Copyright (C) 1993 - 2003 CompuGen Ltd.

CM nucleic - nucleic search, using sw model

Run on: September 22, 2003, 14:02:05 ; Search time 1798.5 Seconds
(without alignments)
610.825 Million cell updates/sec

Title: US-09-475-704A-1
Perfect score: 60
Sequence: 1 gacatcaagcagggccca.....tggacatttcttcaagacc G

Scoring table: IDENTITY_NJC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 452950

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match: 0%
Maximum Match: 100%
Listing first 45 summaries

Database : EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estmu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_hlc:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_hlc:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_esttun:*
- 16: em_estom:*
- 17: em_gss_hum:*
- 18: em_gss_inv:*
- 19: em_gss_pln:*
- 20: em_gss_vrt:*
- 21: em_gss_fun:*
- 22: em_gss_mam:*
- 23: em_gss_mus:*
- 24: em_gss_pro:*
- 25: em_gss_rod:*
- 26: em_gss_phq:*
- 27: em_gss_vrl:*
- 28: gb_gss1:*
- 29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	21.8	36.3	98	28	AF039817
C 2	21.6	36.0	75	29	CC043452
3	21.4	35.7	93	24	W35927
4	20.6	34.3	88	28	BH413628

5	20.6	34.3	88	28	BH413629
6	20.2	33.7	74	9	A1272982
7	20.2	33.7	85	14	R12059
8	20.2	33.7	51	14	T48611
9	20	33.3	74	29	CC043506
10	20	33.3	79	9	AA772825
11	19.8	33.0	81	13	B0876237
12	19.8	33.0	84	14	H27186
13	19.6	32.7	86	9	AA119614
14	19.6	32.7	98	14	H49859
15	19.4	32.3	84	9	AA023110
16	19.2	32.0	70	28	AZ602487
17	19.2	32.0	84	28	BH222409
18	19.2	32.0	85	9	AA237114
19	19.2	32.0	88	9	AA909161
20	19.2	32.0	89	14	H44598
21	19.2	32.0	100	10	BF671124
22	19	31.7	70	9	A1036945
23	19	31.7	70	9	A1154612
24	19	31.7	82	9	AA615572
25	19	31.7	85	9	AA241182
26	18.8	31.3	50	9	A0102354
27	18.8	31.3	55	12	BM285363
28	18.8	31.3	55	12	BM493218
29	18.8	31.3	55	13	B0094077
30	18.8	31.3	64	9	A1020700
31	18.8	31.3	68	12	BF820594
32	18.6	31.0	73	14	H90720
33	18.6	31.0	74	29	CC249304
34	18.6	31.0	79	9	A1117457
35	18.6	31.0	86	9	AA627888
36	18.6	31.0	94	9	AA057469
37	18.6	31.0	99	9	AL829890
38	18.4	30.7	46	28	BH798543
39	18.4	30.7	67	29	CNS020JW
40	18.4	30.7	72	29	CC036825
41	18.4	30.7	83	9	AA117148
42	18.4	30.7	87	28	BH224249
43	18.2	30.3	47	28	BH807575
44	18.2	30.3	61	14	R71917
45	18.2	30.3	85	14	H43578

ALIGNMENTS

RESULT 1	AF039817	Human Homo sapiens	58 bp	DNA	linear	GSS 06-NOV-2000
LOCUS	AF039817	Human Homo sapiens	58 bp	DNA	linear	GSS 06-NOV-2000
DEFINITION	AF039817	Human Homo sapiens	58 bp	DNA	linear	GSS 06-NOV-2000
ACCESSION	AF039817	Human Homo sapiens	58 bp	DNA	linear	GSS 06-NOV-2000
VERSION	AF039817.1	Human Homo sapiens	58 bp	DNA	linear	GSS 06-NOV-2000
KEYWORDS	GSS.	Human Homo sapiens	58 bp	DNA	linear	GSS 06-NOV-2000
SOURCE	Human Homo sapiens	Human Homo sapiens	58 bp	DNA	linear	GSS 06-NOV-2000
ORGANISM	Human Homo sapiens	Human Homo sapiens	58 bp	DNA	linear	GSS 06-NOV-2000
REFERENCE	Whitmore, S.A., Crawford, J., Apostolou, S., Eyre, H., Baker, E., Lower, K.M., Sattasatian, C., Goldup, S., Seshadri, R., Gibson, R.A., Mathew, C.G., Cleton-Janssen, A.M., Savoia, A., Pronk, J.C., Auerbach, A.D., Doggett, N.A., Sutherland, G.R. and Callen, D.F.	Human Homo sapiens	58 bp	DNA	linear	GSS 06-NOV-2000
AUTHORS	Whitmore, S.A., Crawford, J., Apostolou, S., Eyre, H., Baker, E., Lower, K.M., Sattasatian, C., Goldup, S., Seshadri, R., Gibson, R.A., Mathew, C.G., Cleton-Janssen, A.M., Savoia, A., Pronk, J.C., Auerbach, A.D., Doggett, N.A., Sutherland, G.R. and Callen, D.F.	Human Homo sapiens	58 bp	DNA	linear	GSS 06-NOV-2000
TITLE	Construction of a high-resolution physical and transcription map of chromosome 16q24.3: A region of frequent loss of heterozygosity in sporadic breast cancer	Human Homo sapiens	58 bp	DNA	linear	GSS 06-NOV-2000
JOURNAL	Genomics 50 (1), 1-8 (1998)	Human Homo sapiens	58 bp	DNA	linear	GSS 06-NOV-2000
MEDLINE	98292488	Human Homo sapiens	58 bp	DNA	linear	GSS 06-NOV-2000
PUBMED	9628816	Human Homo sapiens	58 bp	DNA	linear	GSS 06-NOV-2000
COMMENT	Contact: Whitmore SA Department of Cytoogenetics and Molecular Genetics Women and Childrens Hospital 72 King William Rd, North Adelaide, SA 5006, Australia	Human Homo sapiens	58 bp	DNA	linear	GSS 06-NOV-2000


```

RESULT 4
BH413628
LOCUS
DEFINITION 1007033F04.1EL_x1 1007 - RescueMu Grid H Zea mays genomic, genomic
survey sequence.
ACCESSION BH413628
VERSION BH413628.1 GI:17589985
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 88)
AUTHORS Walbot,V.
TITLE Maize genomic sequences found using engineered RescueMu transposon
JOURNAL Unpublished
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Very probable ligation site of ends cut by single endonuclease.
Reverse complemented post-ligation sequence from source sequence.
Plate: 1007033 column: 6
Class: transposon-tagged.
FEATURES
source
1..88
Location/Qualifiers
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="mixed background W23/A188/B73"
/db_xref="taxon:4577"
/tissue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="1007 - RescueMu Grid H"
/note="Organ: leaf; Vector: RescueMu (engineered from
pBluescript backbone); Site_1: BamHI; Site_2: BglII;
RescueMu is a 4.9 kb, modified maize Mu transposon
designed to allow plasmid rescue from total genomic DNA.
Mu elements insert preferentially into transcription
units. For more information on RescueMu, go to the web
site 'www.zmdb.iastate.edu' and follow the links for
'RescueMu.' Grid H was grown at Berkeley in 2001. DNA
was extracted from leaf punches, double digested using
BamHI and BglII, and ligated to form circular plasmids.
DH10B cells were transformed and then screened on LB
plates with ampicillin."
BASE COUNT 17 a 29 c 32 g 10 t
ORIGIN
Query Match 34.3%; Score 20.6; DB 28; Length 88;
Best Local Similarity 62.7%; Pred. No. 1.2e+04;
Matches 32; Conservative 0; Mismatches 19; Indels 0; Gaps 0
QY 1 GACATCAGCAGGCGCCCAAGAGGAGCCCTCCCGGAGTACGTGGACCGCTTC 51
||||| 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 38 GACATCTTCGGGACGGGAGGAGGCGCCCTCCCGGAGTACGTGGACCGGATC 88

RESULT 5
BH413629
LOCUS
DEFINITION 1007033F04.y1 1007 - RescueMu Grid H Zea mays genomic, genomic
survey sequence.
ACCESSION BH413629
VERSION BH413629.1 GI:17589989
KEYWORDS GSS.
SOURCE Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 88)
AUTHORS Walbot,V.
TITLE Maize genomic sequences found using engineered RescueMu transposon
JOURNAL Unpublished
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Very probable ligation site of ends cut by single endonuclease.
Reverse complemented post-ligation sequence from source sequence.
Plate: 1007033 column: 6
Class: transposon-tagged.
FEATURES
source
1..88
Location/Qualifiers
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="mixed background W23/A188/B73"
/db_xref="taxon:4577"
/tissue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="1007 - RescueMu Grid H"
/note="Organ: leaf; Vector: RescueMu (engineered from
pBluescript backbone); Site_1: BamHI; Site_2: BglII;
RescueMu is a 4.9 kb, modified maize Mu transposon
designed to allow plasmid rescue from total genomic DNA.
Mu elements insert preferentially into transcription
units. For more information on RescueMu, go to the web
site 'www.zmdb.iastate.edu' and follow the links for
'RescueMu.' Grid H was grown at Berkeley in 2001. DNA
was extracted from leaf punches, double digested using
BamHI and BglII, and ligated to form circular plasmids.
DH10B cells were transformed and then screened on LB
plates with ampicillin."
BASE COUNT 17 a 29 c 32 g 10 t
ORIGIN
Query Match 34.3%; Score 20.6; DB 28; Length 88;
Best Local Similarity 62.7%; Pred. No. 1.2e+04;
Matches 32; Conservative 0; Mismatches 19; Indels 0; Gaps 0
QY 1 GACATCAGCAGGCGCCCAAGAGGAGCCCTCCCGGAGTACGTGGACCGCTTC 51
||||| 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 38 GACATCTTCGGGACGGGAGGAGGCGCCCTCCCGGAGTACGTGGACCGGATC 88

RESULT 5
BH413629
LOCUS
DEFINITION 1007033F04.y1 1007 - RescueMu Grid H Zea mays genomic, genomic
survey sequence.
ACCESSION BH413629
VERSION BH413629.1 GI:17589989
KEYWORDS GSS.
SOURCE Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 88)
AUTHORS Walbot,V.
TITLE Maize genomic sequences found using engineered RescueMu transposon
JOURNAL Unpublished
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Very probable ligation site of ends cut by single endonuclease.
Reverse complemented post-ligation sequence from source sequence.
Plate: 1007033 column: 6
Class: transposon-tagged.
FEATURES
source
1..88
Location/Qualifiers
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="mixed background W23/A188/B73"
/db_xref="taxon:4577"
/tissue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="1007 - RescueMu Grid H"
/note="Organ: leaf; Vector: RescueMu (engineered from
pBluescript backbone); Site_1: BamHI; Site_2: BglII;
RescueMu is a 4.9 kb, modified maize Mu transposon
designed to allow plasmid rescue from total genomic DNA.
Mu elements insert preferentially into transcription
units. For more information on RescueMu, go to the web
site 'www.zmdb.iastate.edu' and follow the links for
'RescueMu.' Grid H was grown at Berkeley in 2001. DNA
was extracted from leaf punches, double digested using
BamHI and BglII, and ligated to form circular plasmids.
DH10B cells were transformed and then screened on LB
plates with ampicillin."
BASE COUNT 17 a 29 c 32 g 10 t
ORIGIN
Query Match 34.3%; Score 20.6; DB 28; Length 88;
Best Local Similarity 62.7%; Pred. No. 1.2e+04;
Matches 32; Conservative 0; Mismatches 19; Indels 0; Gaps 0
QY 1 GACATCAGCAGGCGCCCAAGAGGAGCCCTCCCGGAGTACGTGGACCGCTTC 51
||||| 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 38 GACATCTTCGGGACGGGAGGAGGCGCCCTCCCGGAGTACGTGGACCGGATC 88

```

ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 88)
Walbot,V.
Maize genomic sequences found using engineered RescueMu transposon
Unpublished
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu

Very probable ligation site found so sequence was trimmed.
Post-ligation sequence submitted separately.
Plate: 1007033 column: 6
Class: transposon-tagged.

FEATURES
source

Location/Qualifiers
1..88

/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="mixed background W23/A188/B73"
/db_xref="taxon:4577"
/tissue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"

/clone_lib="1007 - RescueMu Grid H"
/note="Organ: leaf; Vector: RescueMu (engineered from
pBluescript backbone); Site_1: BamHI; Site_2: BglII;
RescueMu is a 4.9 kb, modified maize Mu transposon
designed to allow plasmid rescue from total genomic DNA.
Mu elements insert preferentially into transcription
units. For more information on RescueMu, go to the web
site 'www.zmdb.iastate.edu' and follow the links for
'RescueMu.' Grid H was grown at Berkeley in 2001. DNA
was extracted from leaf punches, double digested using
BamHI and BglII, and ligated to form circular plasmids.
DH10B cells were transformed and then screened on LB
plates with ampicillin."

BASE COUNT 17 a 28 c 33 g 10 t
ORIGIN

Query Match 34.3%; Score 20.6; DB 28; Length 88;
Best Local Similarity 62.7%; Pred. No. 1.2e+04;
Matches 32; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1 GACATCAGCAGGCGCCCAAGAGGAGCCCTCCCGGAGTACGTGGACCGCTTC 51
||||| 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 38 GACATCTTCGGGACGGGAGGAGGCGCCCTCCCGGAGTACGTGGACCGGATC 88

RESULT 5
A1272982

LOCUS
DEFINITION qv63007.x1 NCI_CGAP_Ut1 Homo sapiens cDNA clone IMAGE:1986252 3',
mRNA sequence.

ACCESSION A1272982
VERSION A1272982.1 GI:3895250
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
i (bases 1 to 74)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskalkuk, M.D., Ph.D., Michael R.

MGI:350305

Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Seq primer: -28ML3 rev2 from Amer sham
High quality sequence stop: 1.

FEATURES

Source

FEATURES

Source

Location/Qualifiers
1..88
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:575657"
/sex="male"
/tissue_type="Thymus"
/dev_stage="3 weeks"
/lab_host="DH10R"
/clone_lib="Soares_thymus_2NMT"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTACCAATCTGAAGTGGAGCGCGCTTTTCTTTTCTTTTCTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. RNA provided by Dr. Bertrand Jordan. Library went through two rounds of normalization, and was constructed by Berto Soares and M.Patima Bonaldo."
23 a 11 c 34 g 20 t

BASE COUNT

ORIGIN

Query Match 32.7%; Score 19.6; DB 9; Length 88;

Best Local Similarity 58.6%; Pred. No. 2.5e-04;
Matches 34; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY

3 CATCAAGCAGGCCCCCAAGGAGCCCTTCGGCACTACGTGGACCGCTTCTTCAAGACC 60

Db

74 CCTCAAGCAGCATCCCATGACCTCTTCACCATCTACATCGCAGGCAACACCATTTCC 17

RESULT 14

H49859/c

LOCUS

Y024cl0.s1 Soares adult brain N2b5HB55Y Homo sapiens cDNA clone
IMAGE:178890.3' similar to gb:X17206.40S RIBOSOMAL PROTEIN S4
(HUMAN);, mRNA sequence.

ACCESSION

H49859

VERSION

H49859.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 802
High quality sequence starts: 1
High quality sequence stops: 1
Source: IMAGE Consortium, LLNL.
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Insert Length: 802 Std Error: 0.00

Seq primer: Promega -2lml3

High quality sequence stop: 1.

Location/Qualifiers

1..98

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="GDB:3841086"

/db_xref="taxon:9606"

/clone="IMAGE:178890"

/sex="Male"

/dev_stage="55-year old"

/lab_host="DH10B (ampicillin resistant)"

/clone_lib="Soares adult brain N2b5HB55Y"

/note="Organ: brain; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTACCAATCTGAAGTGGAGCGCGCTTTTCTTTTCTTTTCTTTT 3']; double-stranded cDNA was size selected, ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 53. Library constructed by Berto Soares and M.Patima Bonaldo. The adult brain RNA was provided by Dr. Donald H. Gilden. Tissue was acquired 17-18 hours after death which occurred in consequence of a ruptured aortic aneurysm. RNA was prepared from a pool of tissues representing the following areas of the brain: frontal, parietal, temporal and occipital cortex from the left and right hemispheres, subcortical white matter, basal ganglia, thalamus, cerebellum, midbrain, pons and medulla."

BASE COUNT

ORIGIN

Query Match

32.7%; Score 19.6; DB 14; Length 98;

Best Local Similarity 60.9%; Pred. No. 2.6e+04;

Matches 28; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY

15 CCCAAGGAGCCCTTCGGCACTACGTGGACCGCTTCTTCAAGACC 60

Db

50 CNGCAAGTCCTCCTATCNGGAGTCACTGACCACCTCGTNAAGNCC 5

RESULT 15

AA023110/c

LOCUS

mh66cl0.r1 Soares mouse placenta 4NbMP13.5 14.5 Mus musculus cDNA
clone IMAGE:455922.5', mRNA sequence.

DEFINITION

AA023110

ACCESSION

AA023110

VERSION

AA023110.1

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished
Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:272810

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2003, 14:02:05 ; Search time 193 Seconds
(without alignments)
839.204 Million cell updates/sec

Title: US-09-475-704A-1

Perfect score: 60

Sequence: ; gacatcaagcagggcacaa.....tgaacgcgtttctcaagacc 60

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1345719017 residues

Total number of hits satisfying chosen parameters: 2722626

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03:*

1: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
2: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
3: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
4: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
5: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
6: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
7: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
8: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
9: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
10: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
11: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
12: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
13: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
14: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
15: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
16: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
17: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
18: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
19: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
20: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
21: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA2002.DAT:*
25: /SIDSL/qcgdata/geneseq/geneseq-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	60	100.0	60	21	AAA51607 HIV Gag major homology region nucleotides 844-903.
2	60	100.0	60	24	ABL39974 Synthetic Gag poly
3	55.2	92.0	60	21	AAA51608 HIV Gag major homo
4	53.6	89.3	60	21	AAA70426 HIV p55 Gag Major
5	37.6	62.7	60	21	AAA51627 Major homology reg
6	36	60.0	60	24	ABL39975 Wild type 8_5_TV1
7	34	56.7	90	24	ABK36257 HIV DNA encoding G
8	33	55.0	90	24	ABK36256 HIV DNA encoding G

9	31.8	53.0	87	7	AAAC0663 HIV virus gag regi
10	21	35.0	87	22	ABA71782 Human foetal liver
11	21	35.0	87	22	AAK20151 Human brain expres
12	21	35.0	87	22	AAK46211 Human bone marrow
13	21	35.0	87	22	AA52117 Probe #20803 used
14	21	35.0	87	23	ABS45930 Human liver single
15	21	35.0	87	24	ABS20522 Human genome-deriv
16	20.4	34.0	60	24	AAK32141 CTLB36.1 antigen p
17	20.2	33.7	51	21	AAK76375 Human calmodulin-1
18	20	33.3	81	24	ABS66623 Loop 1b oligonucle
19	19.4	32.3	45	22	AAO09440 HIV reverse trans
20	19.4	32.3	78	24	ABS66624 Loop 1c oligonucle
21	19.2	32.0	51	22	AAK74875 Human silent SNP c
22	19.2	32.0	97	22	AAK76810 Codon-optimised pp
23	19	31.7	53	18	AAK71251 Basic fibroblast g
24	19	31.7	60	24	ABN44208 Human spliced tran
25	19	31.7	65	18	AAK71250 Basic fibroblast g
26	19.8	31.3	30	21	AAA46684 Primer R2 used to
27	18.8	31.3	64	24	ABO75855 Vector pSG1A fragm
28	18.8	31.3	65	24	ABN51784 Mouse spliced tran
29	18.8	31.3	66	23	AAK48738 Pseudomonas aerugi
30	18.6	31.0	47	23	AAH88663 CNS disorder-relat
31	18.6	31.0	51	21	AAK76374 Human calmodulin-1
32	18.6	31.0	51	22	AAK79432 Human silent SNP c
33	18.6	31.0	60	24	ABN36396 Human spliced tran
34	18.6	31.0	65	24	ABN52351 Mouse spliced tran
35	18.6	31.0	65	24	ABN57741 Mouse spliced tran
36	18.4	30.7	37	13	AAQ21583 Sequence of primer
37	18.4	30.7	41	24	AAK04325 Human zinc finger
38	18.4	30.7	51	22	AAK33101 Human SNP oligonuc
39	18.4	30.7	62	10	AAK91755 DNA sequence which
40	18.4	30.7	70	24	ABV95017 Human pancreatic c
41	18.4	30.7	74	22	AAK98662 Human ovarian can
42	18.2	30.3	52	24	ABK88871 Topoisomerase vect
43	18.2	30.3	60	24	ABN49828 Human spliced tran
44	18	30.0	51	22	AAK27470 Human SNP oligonuc
45	18	30.0	60	24	ABN41321 Human spliced tran

ALIGNMENTS

RESULT 1
AAA51607
ID AAA51607 standard; DNA; 60 BP.
XX
AC AAA51607;
XX
CT 31-OCT-2000 (first entry)
XX
DE HIV Gag major homology region nucleotides 844-903.
XX
KW Gag: expression cassette; antigenic; type C; HIV; Env; synthetic;
KW DNA immunization: packaging cell line; antigen presentation; ss.
XX
OS Human immunodeficiency virus.
XX
PN WC2000039304-A2.
XX
PD 06-JUL-2000.
XX
PF 30-DEC-1999; 99WO-US31273.
XX
PR 31-DEC-1998; 98US-0114495.
PR 01-SEP-1999; 99US-0152195.
XX
PA (CHIR) CHIRON CORP.
XX
FI Barnett S. Zur Megede C;
XX
DR WPI: 2000-452401/39.
XX
PI Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV

PT Env polypeptide and the polypeptide useful for immunizing a mammal
PT especially human against HIV
XX
PS Claim 1: Page 92; 113pp; English.
XX
CC Expression cassettes comprising a polynucleotide encoding antigenic
CC type C human immunodeficiency virus (HIV) Gag or Env polypeptides are
CC useful in DNA immunization, generation of packaging cell lines and
CC production of Gag- and/or Env-containing proteins. Synthetic Env and Gag
CC expression cassettes exhibit increased potency for induction of
CC cytotoxic T-lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1
CC self-assemble into non-infectious virus-like particles which are used as
CC a matrix for the proper presentation of an antigen entrapped or
CC associated to the immune system of the host.
XX
SQ Sequence 60 BP; 14 A; 22 C; 15 G; 9 T; 0 other;

Query Match 100.0%; Score 60; DB 21; Length 60;
Best Local Similarity 100.0%; Pred. No. 6.2e-10;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATCAAGCAGGCGCCCAAGGAGCCCTTCGCGACTACGTGGACCGCTTCTTCAAGACC 60
Db 1 GACATCAAGCAGGCGCCCAAGGAGCCCTTCGCGACTACGTGGACCGCTTCTTCAAGACC 60

RESULT 2
ABL39974
ID ABL39974 standard; DNA: 60 BP.
XX
AC ABL39974;
XX
DT 15-MAY-2002 (first entry)
XX
DE Synthetic Gag polynucleotide sequence SEQ ID NO:53.
XX
KW Human immunodeficiency virus type C; antigenic HIV type C proteins;
KW immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
KW immunostimulant; gene therapy; gene; ds.
XX
OS Human immunodeficiency virus type C.
OS Synthetic.
XX
PN WO200204493-A2.
XX
PD 17-JAN-2002.
XX
PF 05-JUL-2001; 2001WO-US21241.
XX
PR 05-JUL-2000; 2000US-0610313.
XX
PA (CHIR) CHIRON CORP.
PA (UYST-) UNIV STELLENBOSCH.
XX
PI Zur Megede J., Barnett SW, Engelbrecht S, Van Rensburg EJ;
XX WPI: 2002-154920/20.
XX
PT New polynucleotides encoding antigenic HIV Type C polypeptides, useful
PT in applications including DNA immunization or generation of packaging
PT cell lines, particularly in gene therapy -
XX
PS Example 1; Fig 24; 233pp; English.
XX
CC The present invention describes expression cassettes comprising a
CC polynucleotide sequence encoding a polypeptide comprising immunogenic
CC HIV type C polypeptides. The expression cassettes comprise any of the
CC HIV type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or
CC Nef (i). (i) have immunostimulant activity and can be used in gene
CC therapy. The HIV type C polynucleotides are useful in applications
CC including DNA immunisation, generation of packaging cell lines, and
CC production of HIV Type C proteins. The polynucleotides are particularly
CC useful in gene therapy and DNA immunisation applications. ABL39942 to

CC ABL40054 and ABL06204 to ABL06215 represent sequences used in the
CC exemplification of the present invention.
XX
SQ Sequence 60 BP; 14 A; 22 C; 15 G; 9 T; 0 other;

Query Match 100.0%; Score 60; DB 24; Length 60;
Best Local Similarity 100.0%; Pred. No. 6.2e-10;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATCAAGCAGGCGCCCAAGGAGCCCTTCGCGACTACGTGGACCGCTTCTTCAAGACC 60
Db 1 GACATCAAGCAGGCGCCCAAGGAGCCCTTCGCGACTACGTGGACCGCTTCTTCAAGACC 60

RESULT 3
AAA51608
ID AAA51608 standard; DNA: 60 BP.
XX
AC AAA51608;
XX
DT 31-OCT-2000 (first entry)
XX
DE HIV Gag major homology region nucleotides 841-900.
XX
KW Gag; expression cassette; antigenic; type C; HIV; Env; synthetic;
KW DNA immunization; packaging cell line; antigen presentation; ss.
XX
OS Human immunodeficiency virus.
XX WO200039304-A2.
XX CG-JUL-2000.
XX PF 30-DEC-1999; 99WO-US31273.
XX PR 31-DEC-1998; 98US-0114495.
XX PR 01-SEP-1999; 99US-0152195.
XX PA (CHIR) CHIRON CORP.
XX Barnett S. Zur Megede J;
PI WPI: 2000-452401/39.
XX
PT Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV
PT Env polypeptide and the polypeptide useful for immunizing a mammal
PT especially human against HIV
XX
PS Claim 1; Page 92; 113pp; English.
XX
CC Expression cassettes comprising a polynucleotide encoding antigenic
CC type C human immunodeficiency virus (HIV) Gag or Env polypeptides are
CC useful in DNA immunization, generation of packaging cell lines and
CC production of Gag- and/or Env-containing proteins. Synthetic Env and Gag
CC expression cassettes exhibit increased potency for induction of
CC cytotoxic T-lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1
CC self-assemble into non-infectious virus-like particles which are used as
CC a matrix for the proper presentation of an antigen entrapped or
CC associated to the immune system of the host.
XX
SQ Sequence 60 BP; 12 A; 24 C; 15 G; 9 T; 0 other;

Query Match 92.0%; Score 55.2; DB 21; Length 60;
Best Local Similarity 95.0%; Pred. No. 2e-08;
Matches 57; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GACATCAAGCAGGCGCCCAAGGAGCCCTTCGCGACTACGTGGACCGCTTCTTCAAGACC 60
Db 1 GACATCCGCGCAGGCGCCCAAGGAGCCCTTCGCGACTACGTGGACCGCTTCTTCAAGACC 60

RESULT 4
AAA70426

```
ID XX AAA70426 standard; DNA; 60 BP.
AC XX AAA70426;
XX XX
DE XX HIV p55 Gag Major Homology Region coding sequence.
XX XX
KW XX HIV-1: AIDS; Gag; vaccine; expression cassette; ss.
XX XX
OS XX Human immunodeficiency virus type 1.
XX XX
PE XX location/Qualifiers
FT CDS 1..60
FT /*tag= a
FT /partial
FT /product= "HIV p55 Gag Major Homology Region protein"
FT /note= "No stop codon given"
XX XX
PN WO200039302-A2.
XX XX
PD 06-JUL-2000.
XX XX
PF 30-DEC-1999; 99WO-US31245.
XX XX
PR 31-DEC-1998; 98US-0114495.
PR 01-DEC-1999; 99US-0158471.
XX XX
PA (CHIR ) CHIRON CORP.
XX XX
PI Barnett S, Zur Megede J, Srivastava I, Lian Y, Hartog K, Liu H;
PI Greer C, Selby M, Walker C;
XX XX
DR WPI: 2000-452400/39.
DR P-PSDB: AAB14214.
XX XX
PT Expression cassettes encoding the human immunodeficiency virus (HIV)
PT Gag-containing polypeptide useful for vaccinating against HIV
PT infections and acquired immunodeficiency syndrome (AIDS) .
XX XX
PS Claim 1: Page 345; 391pp; English.
XX XX
CC The present invention relates to synthetic HIV Gag expression cassettes.
CC The Gag protein of HIV is needed for the assembly of virus-like
CC particles. In addition, the Gag protein is involved in many stages of the
CC HIV life cycle, including assembly, virion maturation after particle
CC release and early post-entry steps in viral replication. The expression
CC cassette may be used for the recombinant expression of HIV
CC Gag-polypeptides which may then be used to vaccinate against HIV
CC infection and acquired immunodeficiency syndrome (AIDS). The present
CC sequence is the coding sequence of HIV p55 Gag Major Homology Region.
CC This sequence is located within the p24-CA sequence of Gag. Mutations in
CC this region can severely impair particle production.
XX XX
SQ Sequence 60 BP; 13 A; 24 C; 15 G; 8 T; 0 other;
XX XX
Query Match 89.3%; Score 53.6; DB 21; Length 60;
Best Local Similarity 93.3%; Pred. No. 6.6e-08;
Matches 56; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 GACATCAAGCAGGGCCCCCAAGGAGCCCTTCCGGGACTACGTGGACCGCTTCTTCAAGACC 60
Db 1 GACATCGCCAGGGCCCCCAAGGAGCCCTTCCGGGACTACGTGGACCGCTTCTTCAAGACC 60
XX XX
RESULT 5
AAA51627
ID AAA51627 standard; DNA; 60 BP.
XX XX
AC AAA51627;
XX XX
DT 31-OCT-2000 (first entry)
XX XX
```

```
DE XX Major homology region of HIV-1 p55 (Gag) coding sequence.
XX XX
KW Gag; expression cassette; antigenic; type C; HIV; Env; synthetic;
KW DNA immunization; packaging cell line; antigen presentation; ss.
XX XX
OS Human immunodeficiency virus type C strain AF110965.
XX XX
PN WO200039304-A2.
XX XX
PD 06-JUL-2000.
XX XX
PF 30-DEC-1999; 99WO-US31273.
XX XX
PR 31-DEC-1998; 98US-0114495.
PR 01-SEP-1999; 99US-0152195.
XX XX
PA (CHIR ) CHIRON CORP.
XX XX
PI Barnett S, Zur Megede J;
XX XX
DR WPI: 2000-452401/39.
DR P-PSDB: AAY95947.
XX XX
PT Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV
PT Env polypeptide and the polypeptide useful for immunizing a mammal
PT especially human against HIV
XX XX
PS Example 1: Page 112; 113pp; English.
XX XX
CC Expression cassettes comprising a polynucleotide encoding antigenic
CC Type C human immunodeficiency virus (HIV) Gag or Env polypeptides are
CC useful in DNA immunization, generation of packaging cell lines and
CC production of Gag- and/or Env-containing proteins. Synthetic Env and Gag
CC expression cassettes exhibit increased potency for induction of
CC cytotoxic T-lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1
CC self-assemble into non-infectious virus-like particles which are used as
CC a matrix for the proper presentation of an antigen entrapped or
CC associated to the immune system of the host.
XX XX
SQ Sequence 60 BP; 24 A; 13 C; 11 G; 12 T; 0 other;
XX XX
Query Match 62.7%; Score 37.6; DB 21; Length 60;
Best Local Similarity 76.7%; Pred. No. 0.0076;
Matches 46; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 1 GACATCAAGCAGGGCCCCCAAGGAGCCCTTCCGGGACTACGTGGACCGCTTCTTCAAGACC 60
Db 1 GACATAAAACAGGACCAACCAAGAGCCCTTTAGAGACTATGTAGACCGGTTCTTIAAAACC 60
XX XX
RESULT 6
ABL39975
ID ABL39975 standard; DNA; 60 BP.
XX XX
AC ABL39975;
XX XX
DT 15-MAY-2002 (first entry)
XX XX
DE Wild type 8_5_Tvl_C.2A Gag major homology region DNA SEQ ID NO:54.
XX XX
KW Human immunodeficiency virus type C; antigenic HIV type C protein;
KW immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
KW immunostimulant; gene therapy; gene; ds.
XX XX
OS Human immunodeficiency virus type C.
XX XX
PN WO200204493-A2.
XX XX
PD 17-JAN-2002.
XX XX
PF 05-JUL-2001; 2001WO-US21241.
XX XX
PR 05-JUL-2000; 2000US-0610313.
```

```
XX (CHIR ) CHIRON CORP.
PA (UYST-) UNIV STELLENBOSCH.
XX
XX Zur Megede J., Barnett SW, Engelbrecht S, Van Rensburg RJ;
XX WPI; 2002-154920/20.
XX
XX New polynucleotides encoding antigenic HIV Type C polypeptides, useful
XX in applications including DNA immunization or generation of packaging
XX cell lines, particularly in gene therapy .
XX
XX Example 1; Fig 25; 233pp; English.
XX
XX The present invention describes expression cassettes comprising a
XX polynucleotide sequence encoding a polypeptide comprising immunogenic
XX HIV type C polypeptides. The expression cassettes comprise any of the
XX HIV type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or
XX Nef (i). (I) have immunostimulant activity and can be used in gene
XX therapy. The HIV type C polynucleotides are useful in applications
XX including DNA immunisation, generation of packaging cell lines, and
XX use of HIV type C proteins. The polynucleotides are particularly
XX useful in gene therapy and DNA immunisation applications. ABL39942 to
XX ABL40054 and ABB06204 to ABB06215 represent sequences used in the
XX exemplification of the present invention.
XX
XX Sequence 60 BP; 24 A; 13 C; 11 G; 12 T; 0 other;
XX
XX Query Match 60.0%; Score 36; DB 24; Length 60;
XX Best Local Similarity 75.0%; Pred. No. 0.024;
XX Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
XX
XX 1 GACATCAAGCAGGCCCCCAAGGAGCCCTTCGCGACTACGTGGACCGCTTCTCAAGACC 60
XX ||||| || || || || || || || || || || || || || || || || || ||
XX 1 GACATAAAACACAGGCCCAAGAGAACCTTTAGAGACTATGAGACCGGTTCTTTAAATC 60
XX
XX
XX RESULT 7
XX ABK36257
XX ID ABK36257 standard; DNA; 90 BP.
XX
XX AC ABK36257;
XX
XX DT 08-MAY-2002 (first entry)
XX
XX DE HIV DNA encoding GAG segment 20.
XX
XX KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;
XX viral infection; human immunodeficiency virus; melanoma;
XX KW bacterial infection; Salmonella; Legionella; parasitic infection;
XX KW Trypanosoma; Toxoplasma; Giardia; ds.
XX
XX OS Human immunodeficiency virus type 1.
XX OS Synthetic.
XX
XX PN WO200190197-A1.
XX
XX PD 29-NOV-2001.
XX
XX PF 25-MAY-2001; 2001WO-AUG0622.
XX
XX PR 26-MAY-2000; 2000AU-0007761.
XX
XX PA (AUSU ) UNIV AUSTRALIAN NAT.
XX
XX PI Thomson SA, Ramshaw IA;
XX
XX WPI; 2002-147575/19.
XX P-PSDB; AAU84418.
XX
XX New synthetic polypeptides having several different segments of at
XX least one parent polypeptide linked together differently compared to
XX the linkage in the parent polypeptide, for inducing immune response
XX
```

```
PT against a pathogen or cancer -
XX
XX Example 1; Fig 12; 364pp; English.
XX
XX The invention relates to a new synthetic polypeptide (I) comprising
XX several different segments of at least one parent polypeptide linked
XX together in a different relationship relative to their linkage in the
XX parent polypeptide to impede, abrogate or otherwise alter at least one
XX function associated with the parent polypeptide and for inducing an
XX immune response against a pathogen or cancer. Also included are a
XX synthetic polynucleotide encoding and a computer system for
XX designing the synthetic polypeptides. The synthetic polypeptides and
XX polynucleotides are referred to as a Savine. The synthetic polypeptide is
XX useful for modulating immune responses preferably directed against a
XX pathogen or a cancer, (e.g., cancers of the lung, breast, ovary, cervix,
XX colon, head and neck, pancreas, prostate, stomach, bladder, kidney, bone
XX liver, oesophagus, brain, testicle, uterus), as potentiating agents.
XX Compositions comprising the polypeptide may be used in the treatment or
XX prophylaxis against viral (such as infections caused by HIV (human
XX immunodeficiency virus), hepatitis, influenza, Japanese encephalitis
XX virus, Epstein-Barr virus and respiratory syncytial virus), bacterial
XX (e.g., infections caused by Neisseria, Meningococcus, Haemophilus,
XX Salmonella, Streptococcus, Legionella and Mycobacterium or parasitic
XX (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,
XX Trypanosoma, Toxoplasma and Giardia) infections. The present
XX sequence encodes a peptide derived from a parent protein used to
XX construct a savine of the invention.
XX
XX Sequence 90 BP; 30 A; 20 C; 22 G; 14 T; 4 other;
XX
XX Query Match 56.7%; Score 34; DB 24; Length 90;
XX Best Local Similarity 74.1%; Pred. No. 0.11;
XX Matches 40; Conservative 2; Mismatches 12; Indels 0; Gaps 0;
XX
XX 7 AAGCAGGGCCCCCAAGGAGCCCTTCGCGACTACGTGGACCGCTTCTCAAGACC 60
XX | : ||||| ||||| ||||| || : ||||| ||||| || : ||||| |||||
XX 1 ARACAGGGGACCCCAAGAGCCCTTTCAGAGACTATGTGGATAGGTTTWCAAAACC 54
XX
XX
XX RESULT 8
XX ABK36256
XX ID ABK36256 standard; DNA; 90 BP.
XX
XX AC ABK36256;
XX
XX DT 08-MAY-2002 (first entry)
XX
XX DE HIV DNA encoding GAG segment 19.
XX
XX KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;
XX viral infection; human immunodeficiency virus; melanoma;
XX KW bacterial infection; Salmonella; Legionella; parasitic infection;
XX KW Trypanosoma; Toxoplasma; Giardia; ds.
XX
XX OS Human immunodeficiency virus type 1.
XX OS Synthetic.
XX
XX PN WO200190197-A1.
XX
XX PD 29-NOV-2001.
XX
XX PF 25-MAY-2001; 2001WO-AUG0622.
XX
XX PR 26-MAY-2000; 2000AU-0007761.
XX
XX PA (AUSU ) UNIV AUSTRALIAN NAT.
XX
XX PI Thomson SA, Ramshaw IA;
XX
XX WPI; 2002-147575/19.
XX P-PSDB; AAU84417.
XX
XX New synthetic polypeptides having several different segments of at
XX
```


CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pat_sequences.
XX
SQ Sequence 87 BP; 17 A; 14 C; 29 G; 27 T; 0 other;
Query Match 35.0%; Score 21; DB 22; Length 87;
Best Local Similarity 73.0%; Pred. No. 1.4e+03;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 20 AGGAGCCCTTCGGCGACTACGTGGACCGCTTCTTCAA 56
||||| ||| ||||| ||||| ||||| |||||
Db 86 AGGAGCCCATCCCAACACTACCTGAACCCACATGATCAA 50
||||| ||| ||||| ||||| ||||| |||||
RESULT 11
AAK20151/c
ID AAK20151 standard; DNA; 87 BP.
XX
AC AAK20151;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe SEQ ID NO: 20142.
XX
DE Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer; ss.
XX
CS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00667.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0508408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
XX
PS Example 4; SEQ ID NO: 20142; 550pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention.
XX
SQ Sequence 87 BP; 17 A; 14 C; 29 G; 27 T; 0 other;
Query Match 35.0%; Score 21; DB 22; Length 87;
Best Local Similarity 73.0%; Pred. No. 1.4e+03;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 20 AGGAGCCCTTCGGCGACTACGTGGACCGCTTCTTCAA 56
||||| ||| ||||| ||||| ||||| |||||
Db 86 AGGAGCCCATCCCAACACTACCTGAACCCACATGATCAA 50
||||| ||| ||||| ||||| ||||| |||||

RESULT 12
AAK46211/c
ID AAK46211 standard; DNA; 87 BP.
XX
AC AAK46211;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed single exon probe SEQ ID NO: 20768.
XX
DE Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukemia; lymphoma; myeloma; ss.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00668.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
PS Example 4; SEQ ID NO: 20768; 658pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukemia and myeloma. The present sequence is one of
CC the probes of the invention.
XX
SQ Sequence 87 BP; 17 A; 14 C; 29 G; 27 T; 0 other;
Query Match 35.0%; Score 21; DB 22; Length 87;
Best Local Similarity 73.0%; Pred. No. 1.4e+03;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 20 AGGAGCCCTTCGGCGACTACGTGGACCGCTTCTTCAA 56
||||| ||| ||||| ||||| ||||| |||||
Db 86 AGGAGCCCATCCCAACACTACCTGAACCCACATGATCAA 50
||||| ||| ||||| ||||| ||||| |||||
RESULT 13
AAI52117/c
ID AAI52117 standard; DNA; 87 BP.
XX
AC AAI52117;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #20803 used to measure gene expression in human placenta sample.
XX
DE Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder; ss.
XX

OS Homo sapiens.
XX WO200157272-A2.
PN
XX
PD
XX
XX 09-AUG-2001.
PF
XX 30-JAN-2001; 2001WO-US000663.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000US-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI: 2001-488897/53.
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human placenta -
XX
PS Claim 25: SEQ ID No 20803; 654pp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders.
XX
SQ Sequence 87 BP; 17 A; 14 C; 29 G; 27 T; 0 other;

Query Match 35.0%; Score 21; DB 22; Length 87;
Best Local Similarity 73.0%; Pred. No. 1.4e+03;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 20 AGGAGCCCTTCGCGACTACGTGGACCGCTTCTTCAA 56
DB 86 AGGAGCCCATCCACACTACTGACCATGATCAA 50

RESULT 14
ABS45930/c
ID ABS45930 standard; DNA; 87 BP.
XX
AC ABS45930;
XX
XX 25-FEB-2003 (first entry)
XX
DE Human liver single exon probe, SEQ ID No 20920.
XX
KW Human; single exon nucleic acid probe; liver; cirrhosis;
KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW coronary heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157273-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000664.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000US-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI: 2001-488898/53.
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human adult liver -
XX
PS Claim 4: SEQ ID No 20920; 658pp; English.
XX
CC The invention relates to a single exon nucleic acid probe (SENP) (1) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult
CC liver. (1) may be used for predicting, measuring and displaying gene
CC expression in samples derived from human adult liver. The genes
CC identified may be involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which
CC is associated with coronary heart disease. ABS25011-ABS51005 represent
CC human liver single exon nucleic acid probes of the invention.
CC Note: The sequence information for this patent does not appear in the
CC printed specification but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 87 BP; 17 A; 14 C; 29 G; 27 T; 0 other;

Query Match 35.0%; Score 21; DB 23; Length 87;
Best Local Similarity 73.0%; Pred. No. 1.4e+03;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 20 AGGAGCCCTTCGCGACTACGTGGACCGCTTCTTCAA 56
DB 86 AGGAGCCCATCCACACTACTGACCATGATCAA 50

RESULT 15
ABS20522/c
ID ABS20522 standard; DNA; 87 BP.
XX
AC ABS20522;
XX
XX 19-AUG-2002 (first entry)
XX
DE Human genome-derived single exon probe ORF from lung SEQ ID No 20513.
XX
KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary syndrome;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease; open reading frame; ORF.
XX
OS Homo sapiens.
XX
PN WO200186003-A2.
XX
PD 15-NOV-2001.
XX
PF 30-JAN-2001; 2001WO-US000665.
XX
PR 04-FEB-2000; 2000US-180312P.
PR 26-MAY-2000; 2000US-207456P.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-234687P.
PR 27-SEP-2000; 2000US-236359P.
PR 04-OCT-2000; 2000US-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.
XX PA
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2002-114183/15.
XX PT Spatially-addressable set of single exon nucleic acid probes, used to
XX PT measure gene expression in human lung samples
XX PS
XX PS Claim 4; SEQ ID No 20513; 634pp; English.
XX CC The invention relates to a spatially-addressable set of single exon
XX CC nucleic acid probes for measuring gene expression in a sample derived
XX CC from human lung comprising single exon nucleic acid probes having one of
XX CC 12614 nucleic acid sequences mentioned in the specification, or their
XX CC complements or the 12387 open reading frames derived from the 12614
XX CC probes. Also included are a microarray comprising the novel set of
XX CC probes; the novel set of probes which hybridize at high stringency to a
XX CC nucleic acid expressed in the human lung; measuring gene expression in a
XX CC sample derived from human lung, comprising (a) contacting the array with
XX CC a collection of detectably labeled nucleic acids derived from human lung
XX CC mRNA, and (b) measuring the label detectably bound to each probe of
XX CC the array; identifying exons in a eukaryotic genome, comprising
XX CC (a) algorithmically predicting at least one exon from genomic sequences
XX CC of the eukaryote; and (b) detecting specific hybridisation of detectably
XX CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
XX CC having a fragment identical to the predicted exon, the probe is included
XX CC in the above mentioned microarray; assigning exons to a single gene,
XX CC comprising (a) identifying exons from genomic sequence by the method
XX CC above and (b) measuring the expression of each of the exons in several
XX CC tissues and/or cell types using hybridisation to a single exon
XX CC microarrays having a probe with the exon, where a common pattern of
XX CC expression of the exons in the tissues and/or cell types indicates that
XX CC the exons should be assigned to a single gene; a peptide comprising one
XX CC of 12011 sequences, mentioned in the specification, or encoded by the
XX CC probes/open reading frames (ORF). The probes are used for gene
XX CC expression analysis, and for identifying exons in a gene, particularly
XX CC using human lung derived mRNA and for the study of lung diseases
XX CC such as asthma, lung cancer, chronic obstructive pulmonary disease
XX CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
XX CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
XX CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
XX CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
XX CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
XX CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
XX CC and hyaline membrane disease. The present sequence is a single exon
XX CC probe open reading frame of the invention.
XX CC Note: The sequence data for this patent did not form part
XX CC of the printed specification, but was obtained in electronic
XX CC format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 87 BP; 17 A; 14 C; 29 G; 27 T; 0 other;

Query Match 35.0%; Score 21; DB 24; Length 87;
Best Local Similarity 73.0%; Pred. No. 1.4e+03;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 20 AGGAGCCCTCCGGGACIACGTGGACCGCTTCTCAA 55
||||| ||| ||||| ||||| ||||| |||||
Db 86 AGGAGCCCATCCACACTACCTGACCATCATGATCAA 50